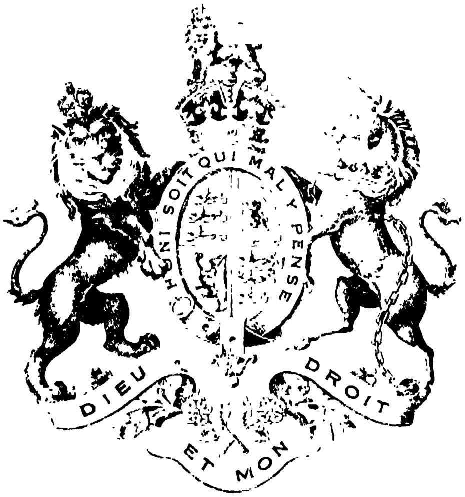

HONI SOIT QUI MAL Y PENSE
DIEU ET MON DROIT

AGRICULTURAL RESEARCH INSTITUTE

PUSA

314709

2.4.23

PROCEEDINGS

OF THE

ROYAL SOCIETY OF LONDON

SERIES A

CONTAINING PAPERS OF A MATHEMATICAL AND PHYSICAL CHARACTER.

VOL. CII.

314709
IARI

LONDON:
PRINTED FOR THE ROYAL SOCIETY AND SOLD BY
HARRISON AND SONS, LTD., ST. MARTIN'S LANE,
PRINTERS IN ORDINARY TO HIS MAJESTY.

MARCH, 1923.

LONDON:
HARRISON AND SONS, LTD., PRINTERS IN ORDINARY TO HIS MAJESTY,
ST. MARTIN'S LANE.

CONTENTS.

SERIES A. VOL. CII.

Minutes of Meetings, November 2, 9, 16, 23, 30; December 7, 1922; January 18, 25; February 1, 8, 15, 22, 1923.

No. A 714.—October 2, 1922.

PAGE

Spontaneous Incandescence of Substances in Atomic Hydrogen Gas. By R. W. Wood, For. Mem. R.S. 1

On the Scattering of β-Rays. By H. A. Wilson, F.R.S., Rice Institute, Houston, Texas, U.S.A. 9

A Study of Catalytic Actions at Solid Surfaces. Part VIII. The Action of Sodium Carbonate in Promoting the Hydrogenation of Phenol. By E. F. Armstrong, D.Sc., F.R.S., and T. P. Hilditch, D.Sc. 21

A Study of Catalytic Actions at Solid Surfaces. Part IX. The Action of Copper in Promoting the Activity of Nickel Catalyst. By E. F. Armstrong, D.Sc., F.R.S., and T. P. Hilditch, D.Sc. 27

On the Absorption of λ 5460·97 A. by Luminous Mercury Vapour. By Prof. J. C. McLennan, F.R.S., D. S. Ainslie, M.A., and Miss F. M. Cale, B.A., University of Toronto. (Plate 1) 33

The Effect of Temperature on Gravitative Attraction. By P. E. Shaw, B.A., D.Sc., and N. Davy, B.Sc., A.R.C.S. Communicated by Prof. E. H. Barton, F.R.S.... 46

The Loss of Energy of an α-Ray Beam in its Passage through Matter. Part I. Passage through Air and CO_2. By P. L. Kapitza, Lecturer in Physics in the Polytechnic Institute, Petrograd. Communicated by Prof. Sir E. Rutherford, F.R.S. 48

The Relation between the Evolution of Heat and the Supply of Energy during the Passage of an Electric Discharge through Hydrogen. By J. Keith Roberts, M.Sc., Trinity College, Cambridge; Exhibition of 1851 Scholar of the University of Melbourne. Communicated by Prof. Sir E. Rutherford, F.R.S. 72

Free Motion of a Sphere in a Rotating Liquid parallel to the Axis of Rotation. By S. F. Grace, University of Liverpool. Communicated by G. I. Taylor, F.R.S. 89

No. A 715.—November 1, 1922.

On the Mutual Threading of Vortex Rings. By W. M. Hicks, F.R.S. 111

A Spectroscopic Investigation of the Ionisation of Argon by Electron Collisions. By Frank Horton, Sc.D., Professor of Physics in the University of London, and Ann Catherine Davies, D.Sc., Royal Holloway College. Communicated by Sir J. J. Thomson, O.M., F.R.S. 131

PAGE

The Molecular Scattering of Light in Vapours and in Liquids and its Relation to the Opalescence observed in the Critical State. By K. R. Ramanathan, M.A., Demonstrator in Physics, H.H. the Maharajah's College, Trivandrum; University of Madras Research Scholar, Calcutta. Communicated by Dr. Gilbert T. Walker, F.R.S. 151

The Motion of Ellipsoidal Particles Immersed in a Viscous Fluid. By G. B. Jeffery, M.A., D.Sc., Fellow of University College, London. Communicated by Prof. L. N. G. Filon, F.R.S. 161

The Motion of a Sphere in a Rotating Liquid. By G. I. Taylor, F.R.S. 180

Polarisation of the Light Scattered by Mercury Vapour near the Resonance Periodicity. By Lord Rayleigh, F.R.S. 190

The Scattering of Hydrogen Positive Rays, and the Existence of a Powerful Field of Force in the Hydrogen Molecule. By G. P. Thomson, M.A., Fellow of Corpus Christi College, Cambridge, Professor of Natural Philosophy in the University of Aberdeen. Communicated by Sir J. J. Thomson, F.R.S. 197

Multenions and Differential Invariants.—II. By Alex. McAulay, Professor of Mathematics in the University of Tasmania. Communicated by W. B. Hardy, Sec.R.S. 210

No. A 716.—December 1, 1922.

On the Characteristics of Cylindrical Journal Lubrication at High Values of the Eccentricity. By T. E. Stanton, D.Sc., F.R.S. 241

On the Fluorescence of Æsculin. By J. C. McLennan, F.R.S., Professor of Physics, and Miss F. M. Cale, M.A., University of Toronto 256

The Propagation of Gravitational Waves. By A. S. Eddington, F.R.S. 268

A New Method for Studying Ionising Potentials. By H. D. Smyth, Ph.D., National Research Fellow in Physics, U.S.A. Communicated by Prof. Sir Ernest Rutherford, F.R.S. 283

On the Analysis of α-Ray Photographs. By P. M. S. Blackett, B.A., Charles Kingsley Bye-Fellow, Magdalene College, Cambridge. Communicated by Prof. Sir E. Rutherford, F.R.S. (Plate 2.) 294

On the Structure and Chemical Activity of Copper Films, and the Colour Changes accompanying their Oxidation. By C. N. Hinshelwood, B.A., Fellow of Trinity College, Oxford. Communicated by Prof. J. W. Nicholson, F.R.S. 318

On Measurements of Electrode Potential Drop with Direct Current and Alternating Current Electrolysis. By S. Marsh, B.Sc., Ph.D., and A. E. Evans, B.Sc. Communicated by E. H. Griffiths, F.R.S. 328

Variation of the Intensity of Reflected X-Radiation with the Temperature of the Crystal. By Ivor Backhurst, B.Sc. Communicated by Sir William Bragg, K.B.E., F.R.S. 340

PAGE

An Investigation of the Colour Vision of 527 Students by the Rayleigh Test. By R. A. Houstoun, D.Sc., Lecturer on Physical Optics in the University of Glasgow. Communicated by Prof. A. Gray, F.R.S. 353

The Line of Action of the Resultant Pressure in Discontinuous Fluid Motion. By S. Brodetsky, M.A., Ph.D., University of Leeds. Communicated by Sir George Greenhill, F.R.S. 361

No. A 717.—January 1, 1923.

Address of the President, Sir Charles S. Sherrington, at the Anniversary Meeting, November 30, 1922 373

On the Excitation of Characteristic X-Rays from Light Elements. By J. C. McLennan, F.R.S., Professor of Physics, and Miss M. L. Clark, B.A., University of Toronto 389

A Theory of Meteors, and the Density and Temperature of the Outer Atmosphere to which it Leads. By F. A. Lindemann, F.R.S., Professor of Experimental Philosophy, and G. M. B. Dobson, M.A., Lecturer in Meteorology, University of Oxford 411

The Theory of the Scattering of α and β-Rays. By J. H. Jeans, Sec.R.S. 437

Spectrum of Active Nitrogen as Affected by Admixture of the Inert Gases, with a Note on the Origin of the Cyanogen Spectrum. By Lord Rayleigh, F.R.S. (Plate 3) 453

Directional Observations of Atmospheric Disturbances, 1920–21. By R. A. Watson Watt, B.Sc., F.Inst.P. Communicated by Sir Henry Jackson, F.R.S. 460

The Quantum Theory and Electromagnetic Phenomena. By William Wilson, Ph.D., D.Sc., Professor of Physics in the University of London. Communicated by Prof. J. W. Nicholson, F.R.S. 478

The Carbon Arc Spectrum in the Extreme Ultra-Violet. By F. Simeon, B.Sc., F.Inst.P., Physicist in the Research Laboratory of Adam Hilger, Ltd. Communicated by Prof. A. W. Porter, F.R.S. (Plate 4) 484

Changes in the Charge of an α-Particle passing through Matter. By G. H. Henderson, Ph.D., 1851 Exhibition Scholar of Dalhousie University, Halifax, N.S. Communicated by Prof. Sir E. Rutherford, F.R.S. (Plate 4) 496

No. A 718.—February 1, 1923.

The Crystalline Structure and Properties of Tartaric Acid. By W. T. Astbury, B.A. Communicated by Sir William Bragg, F.R.S. 506

On the Quantum Theory of the Simple Zeeman Effect. By A. M. Mosharrafa, B.Sc., University of London, King's College. Communicated by Prof. O. W. Richardson, F.R.S. 529

The Magnitude of the Gyromagnetic Ratio. By O. W. Richardson, F.R.S., King's College, London 538

PAGE

Stability of a Viscous Liquid contained between Two Rotating Cylinders. By G. I. Taylor, F.R.S. (Abstract) 541

Discontinuous Fluid Motion Past Circular and Elliptic Cylinders. By S. Brodetsky, M.A., Ph.D., University, Leeds. Communicated by Prof. G. H. Bryan, F.R.S. 542

The Propagation of Earthquake Waves. By J. H. Jeans, Sec.R.S. 554

The Coincidence Method for the Wave-Length Measurement of Absorption Bands. By H. Hartridge. Communicated by Prof. J. N. Langley, F.R.S. 575

On the Changes in Iron and Steel at Temperatures below 280 C. By Frank Charles Thompson, Professor of Metallurgy, University of Manchester, and Edwin Whitehead. Communicated by Prof. H. C. H. Carpenter, F.R.S. 587

Smokes: Part I.—A Study of their Behaviour and a Method of Determining the Number of Particles they Contain. By R. Whytlaw-Gray, J. B. Speakman, and J. H. P. Campbell. Communicated by Prof. A. Smithells, F.R.S. (Plate 5) 600

Smokes: Part II.—A Method of Determining the Size of the Particles in Smokes. By R. Whytlaw-Gray and J. B. Speakman. Communicated by Prof. A. Smithells, F.R.S. 615

The Significance of the Electrode Potential. By Jaroslav Heyrovský, D.Sc., Ph.D. Communicated by Prof. F. G. Donnan, F.R.S. 628

The Effect of Long Grinding on Quartz (Silver Sand). By Rames C. Ray, M.Sc. Communicated by Dr. M. W. Travers, F.R.S. 640

No. A 719.—March 1, 1923.

Bakerian Lecture: The Distortion of an Aluminium Crystal during a Tensile Test. By G. I. Taylor, F.R.S., and C. F. Elam 643

Studies of Iridescent Colour, and the Structure producing it.—I. The Colours of Potassium Chlorate Crystals. By Lord Rayleigh, F.R.S. (Plates 6 and 7) ... 668

Studies of Iridescent Colour, and the Structure producing it.—II. Mother-of-Pearl. By Lord Rayleigh, F.R.S. (Plate 8) 671

Investigations on the Paramagnetic Sulphates at Low Temperatures. By L. C. Jackson, M.Sc., A.Inst.P., "1851 Exhibition Research Scholar," and Prof. H. Kamerlingh Onnes, For.Mem.R.S. 678

Investigations on the Paramagnetism of Crystals at Low Temperatures. By L. C. Jackson, M.Sc., A.Inst.P., "1851 Exhibition Research Scholar," and Prof. H. Kamerlingh Onnes, For.Mem.R.S. 680

Pleochroic Haloes of Various Geological Ages. By J. Joly, Sc.D., F.T.C.D., F.R.S. (Plate 9) 682

The Potential Difference occurring in a Donnan Equilibrium and the Theory of Colloidal Behaviour. By A. V. Hill, F.R.S. 705

PAGE

The Periodic Opacity of Certain Colloids in Progressively Increasing Concentrations of Electrolytes. By J. Holker, D.Sc., M.D. Communicated by Prof. A. V. Hill, F.R.S. 710

The Kinetic Energy of Electrons Emitted from a Hot Tungsten Filament. By J. H. Jones, B.Sc, Ph.D. Communicated by Prof. O. W. Richardson, F.R.S ... 734

The Production of Artificial Vowel Sounds. By Sir R. A. S. Paget, Bart. Communicated by Sir W. Bragg, F.R.S. 752

On the Steady Motion of a Cylinder through Infinite Viscous Fluid. By Arthur Berry, M.A., Fellow of King's College, Cambridge, and Lorna M. Swain, Fellow of Newnham College, Cambridge. Communicated by H. W. Richmond, F.R.S. 766

Obituary Notices of Fellows Deceased:—

Sir Alfred Bray Kempe (with portrait) i
Sir William Christie xi
William Gowland (with portrait) xvi
E. H. Grove-Hills xx
G. W. Walker xxii
C. G. Knott xxvii
Jacobus Cornelius Kapteyn xxix

Index xxxvii

PROCEEDINGS OF

THE ROYAL SOCIETY.

Section A.—Mathematical and Physical Sciences.

Spontaneous Incandescence of Substances in Atomic Hydrogen Gas.

By R. W. Wood, For. Mem. R.S.

(Received June 26, 1922.)

In a previous communication* it has been shown that if a very long vacuum tube of moderate bore, filled with hydrogen at a pressure of ½ mm., is operated by a direct or alternating high potential current, the secondary spectrum appears only at the ends of the tube in the vicinity of the electrode bulbs, the central portion showing the lines of the Balmer series, with a faint trace only of the secondary spectrum. By this method photographs of the series down to the twentieth member were obtained. In more recent work the series has been photographed to the eighteenth line in the third order spectrum of a 7-inch plane grating with a lens of 20-feet focus, and the wave-lengths determined to within a few thousandths of an Ångström. This work will be described elsewhere. Practically all of the very peculiar effects described in the paper referred to above have been explained, and in the pursuit of some of the more elusive phenomena, some extremely interesting properties of atomic hydrogen gas have come to light which will be described briefly in the present paper.

The work developed out of a study of what I referred to in the earlier paper as "Infected Spots." It was frequently observed that white spots sometimes appeared along the central portion of the tube, which normally is fiery-purple in colour, and almost invisible through a green colour filter. These spots showed the secondary spectrum of hydrogen, with the full intensity exhibited at the ends of the tube, some fifty times as intense as in adjacent portions of the tube which gave the nearly pure Balmer spectrum.

---

* 'Roy. Soc. Proc.,' A, vol. 97 (1920).

Though I suspected that these infected spots, which often appeared quite suddenly and were difficult to get rid of, were in some way due to a contamination of the wall of the tube, I was quite unable to understand how they were produced, or why they caused the appearance of the secondary spectrum.

The accidental entrance into the tube of a minute speck of sealing-wax from the joint leading to the pump started the correct line of attack. Under the action of the powerful discharge the fragment was almost immediately reduced to a white spot of stannous oxide, and the hydrogen secondary spectrum at once appeared in its vicinity. After operating the tube for a few minutes, examination of the spot of oxide with a lens showed that it was covered with minute silvery globules of reduced tin.

I next tried fusing a fine wire of tungsten into the wall of the tube (which was of pyrex glass), fig. 1. The wire was raised to incandescence as soon as the discharge was started, and the secondary spectrum came out strong in its vicinity. This surprised me very much, as it seemed incredible that the wire could be brought to a white heat by any electrical action, and it appeared unlikely that the mere high temperature of the hydrogen was responsible. A thread of soft glass introduced into the tube was not melted, though completely immersed in the discharge. This showed that it was not high temperature of the gas that caused the incandescence of the wire. To test whether or not the secondary spectrum was in any way due to the emission of electrons by the wire and oxide speck, I arranged a small tungsten spiral in a side tube, out of the line of the discharge, and connected it to a well insulated storage battery. The spiral came to a red heat, however, *even when disconnected from the battery.* This indicated clearly that the heating of the wire was, in all probability, due to the action of the surface of the metal in causing re-combination of dissociation products of the discharge.

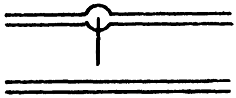


Fig. 1.

A platinum wire was next introduced into the discharge, but it remained non-luminous even with the heaviest current which could be brought to bear. With air at 0·5 mm. in the tube the platinum wire immediately came to a red heat, doubtless by causing the re-combination of nitrogen atoms.

During all of these experiments, pure electrolytic hydrogen entered one end of the tube through a very fine long capillary. No drying agent was employed, as the presence of the water vapour, which comes over with the hydrogen, is necessary for the production of the pure Balmer spectrum. The probable reason of this much discussed circumstance will appear presently.

As to the nature of the surface reaction which caused the heating of the wire, it seemed probable that it was the re-combination of hydrogen atoms to form molecules, as the very small amount of oxygen present could hardly account for the phenomenon.

As was shown in the earlier paper, the secondary spectrum flashes out for a small fraction of a second when the current is first turned on; the duration of this flash was subsequently found* to be from 0·01 second (with 20 ampères in the primary of the transformer) to 0·04 second (with 2 ampères). If the switch is opened and kept open for a time in excess of about 0·2 second, the secondary flash can be obtained again on closing the switch; but if the time of the interruption of the current is much less than this, only the Balmer lines appear when the current is turned on again. The interruptions which occur during the normal operation of the tube, due to the passage of the potential of the transformer through the zero point, are too brief to allow the restoration of the condition necessary for the development of the secondary spectrum flash; in other words, the hydrogen remains in the atomic state. The simplest explanation of this is to consider that under the action of a heavy current practically all of the hydrogen in the central portion of the tube is kept permanently in the atomic condition, and hence only the Balmer lines appear.

Since the secondary spectrum flash does not appear when the current is broken, but only at the make, it seems safe to assume that the secondary spectrum is produced either by the continuous excitation of hydrogen molecules or by the explosion of the molecule into atoms, but not by the recombination of the atoms to form molecules. The great mystery was why the presence of water vapour or oxygen was necessary to suppress the secondary, and bring out the atomic, spectrum (Balmer series). I tried adding increasing amounts of oxygen to the hydrogen stream, by heating a small bulb containing crystals of permanganate of potash, joined to the tube by a short capillary. The tube contained a tungsten and a platinum wire; the former glowed at full incandescence while the latter remained dark. On gradually increasing the amount of oxygen, the tungsten wire cooled off and became black while the platinum wire grew first red then white hot.

The experiment was repeated with hydrogen dried by passage over P_2O_5. Neither wire heated until the oxygen stream was started, when the tungsten came to a white heat, and as more oxygen was added, cooled off again, the platinum heating in its turn as before. On cutting down the supply of oxygen the phenomena occurred in reverse order.

I discussed these results with Dr. Irving Langmuir, who has made an extended study of the atomic hydrogen produced by passing the gas over an incandescent surface of tungsten. He at once pointed out the property which oxygen has of "poisoning" a catalytic body. This explained why the addition of more oxygen stopped the heating of the tungsten, if we assume that the heating of the wire is caused by the catalytic action of its surface in bringing about the re-combination of the hydrogen atoms.

---

* 'Phil. Mag.,' vol. 43, pp. 9, 736 (1921).

At an earlier period in the course of the experiments, when I first speculated on the possible formation of molecular hydrogen by the tungsten wire, I tried introducing the hydrogen through a capillary which opened directly into the discharge at a point near the middle of the tube, where only the Balmer lines appeared, fully expecting the secondary spectrum to appear in tht vicinity of the mouth of the capillary. To my great surprise, the discharge at the point remained fiery purple, with no increase in the intensity of the faint secondary spectrum.

As the hydrogen was flowing in at the rate of about 1 c.c. per minute, it seemed strange that it could be exploded into the atomic condition without giving any visible increase in the amount of secondary spectrum. Dr. Langmuir, however, made a calculation which shows that the wire may generate molecular hydrogen out of atomic at a much greater rate than the rate of inflow.

Assuming that the wire was 1 cm. long and 0·2 mm. in diameter, he found that it would take 4 watts to maintain it at 2,400° K. To supply this energy 0·25 c.c. of hydrogen (at atmospheric pressure) must be produced every second. This corresponds to 15 c.c. in 1 minute, or fifteen times as great as the rate of inflow. He further found that "when 1 grm. of hydrogen atoms combine to form molecular hydrogen, 45,000 small calories of heat are produced. In atomic hydrogen at 1 mm. pressure and 500° C. [my estimate of temperature of the gas in the tube, R.W.W.] the amount of hydrogen that strikes each square centimetre of surface per second is 0·0026 grm. This would produce a heating effect of 490 watts per square centimetre, which would be sufficient to heat a tungsten surface to 4100° absolute. Assuming the wire to be at 2400°, I calculated that this temperature could be maintained by a pressure of 0·16 mm. of atomic hydrogen in the tube, which seems a reasonable value."

It thus appears that the heating of the wire can be fully accounted for, and that the amount of molecular hydrogen generated from atomic is so much greater than the inflow from the capillary, that we need no longer be surprised at the failure of the secondary spectrum to appear at the point where the gas entered the discharge tube.

Langmuir's observation that the presence of minute traces of oxygen in the hydrogen prevents the formation of atomic hydrogen by a tungsten wire raised to incandescence in the gas by an electric current, and his suggestion that the oxygen must act in the same manner when a comparatively cool tungsten wire is causing the recombination of the atoms, and absorbing the liberated heat, is of immense value in clearing up many of the mysteries of the hydrogen tube.

Previous to the experiments with the metallic wires I had found that if the inner surface of tube was fine ground with carborundum at a point near the centre, this portion of the tube remained always in the "white stage," *i.e.*, the secondary spectrum was strong and the Balmer series weak.

The ground surface evidently caused the recombination of the atomic hydrogen as fast as it was formed by the discharge, and the concentration never became great. Just why this was so was not at first apparent. Experiments with a thermocouple in the earlier work had shown that the temperature of the outer wall of the tube was always higher at the "infected spots" than on portions of the tube showing the Balmer lines only. The same was true of the ground-glass tube, the temperature rising to such a point that the D lines of sodium came out so strong that the discharge appeared yellow in spots.

The increase in the amount of glass surface exposed to the gas by the grinding could hardly account for the increased catalytic action, and it occurred to me that a fractured surface might be more active in bringing about the combination of the atoms. I accordingly made a pear-shaped bead of pyrex glass, scratched it across with a glass cutter and cracked it in two. This bead was suspended in the discharge by a glass fibre as shown in fig. 2.

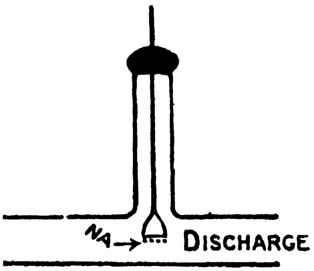


Fig. 2.

The cracked bead caused the discharge to become white in its vicinity while a whole bead could be immersed in the purple discharge without affecting it in the least. Moreover, the flat under-surface of the bead (the clean fracture) was seen to be covered with glowing sodium vapour, when the bead was viewed through a direct vision prism (to "spectroheliograph" it). This showed that the fractured surface had been raised to a very high temperature by the atomic hydrogen, while the fire-polished surface remained comparatively cool.

The action of other metals and oxides was next examined. A strip of very clean thin aluminium foil immersed in the purple discharge caused the appearance of the secondary spectrum in its vicinity, but after a few minutes' operation the white discharge in the vicinity of the narrow strip of foil disappeared, and the aluminium was as neutral to the atomic hydrogen as the fire-polished bead. Operating the tube with air in it restored the aluminium to its original condition, and the white hydrogen discharge was again obtained.

It had frequently been observed, in working with long hydrogen tubes, that sometimes the white secondary spectrum discharge reached to a distance of 30 or 40 cm. from the electrode bulbs, while at other times, with pyrex tubes, carefully cleaned with hot chromic acid, and clean new aluminium electrodes, the purple discharge came up to within 2 or 3 cm. of the bulbs. It now seems probable that the secondary spectrum at the ends of the tube is due to the re-formation of molecular hydrogen by the more or less oxidised electrodes, in other words, the secret of obtaining the extended Balmer series by the use of a long tube lay in the utilisation of a portion of the tube far removed from the catalysing action of the electrodes.

A sufficiently powerful discharge is able to hold practically all of the hydrogen in the atomic condition. The secret of obtaining the complete Balmer series probably lies in obtaining a tube wall of zero catalysing power; I have tried fused quartz, but it appears to be no better than pyrex glass. Next autumn I plan to make a jointed tube of as many different types of glass as possible. It must be remembered, however, that it is the oxygen which brings the glass into the required condition.

A copper wire, strongly oxidised by a flame when introduced into the hydrogen tube, gave a strong secondary spectrum, and the surface immediately became of a bright metallic copper colour. Zinc oxide gave a black deposit on the glass wall in its vicinity. In these last two experiments the substances were in lateral tubes just outside of the discharge.

The most interesting material, however, was thorium oxide, with a trace of cerium (fragments of a Welsbach mantle). Small specks of this substance, visible only with a lens introduced into the tube, were scattered about by the discharge, and, adhering to the walls, glowed like first magnitude stars with a colour which appeared distinctly greenish, in contrast to the deep purple of the hydrogen discharge in which they were immersed. They produced no appreciable amount of secondary spectrum on account of their small size. Larger fragments produced strong secondary spectrum without coming to incandescence. This is to be expected, for if the atomic hydrogen acts upon too large a mass its energy of combination is distributed over too much material to bring it to incandescence. If, for example, we have a tungsten wire in a tube containing specks of thoria, the specks all glow brilliantly when the current is turned on. Presently the wire begins to act on the hydrogen, the secondary spectrum appears, and the incandescent thoria specks in the vicinity of the wire are immediately extinguished.

The ability of the tungsten wire to cause the re-combination of the atomic hydrogen appears to depend to a certain extent on the condition of the surface. It was frequently observed that the wire did not heat until it had been immersed in the discharge for a minute or two, when it suddenly rose to incandescence; not until this change took place did the purple colour of the discharge change to white. After continued operation the wire often cooled down to a dull red heat. If the current was then shut off, and the wire allowed to rest for 5 or 10 minutes in the gas stream, it usually became white hot the moment the discharge was started again.

To investigate the flow of the atomic hydrogen out of the discharge tube the arrangement shown in fig. 3 was used.

Metallic wires, mounted on slender rods of glass, could be introduced through the tube A and placed at varying distances from the discharge, and the flow of the gas to greater distances could be followed by specks of thoria beyond the U-bend in the tube leading to the pump.

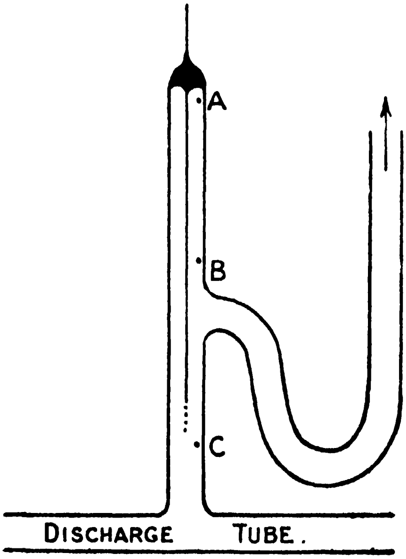


FIG. 3.

It was found that a tungsten wire could be maintained at a red heat at a distance of 4 cm. from the discharge tube (the tube was 5 mm. in diameter), while thoria particles were highly luminous at distances of 20 cm. or more. The increase in the free path of the atoms as the pressure was lowered was beautifully shown in one experiment, in which three thoria particles were attached to the wall at A, B, and C. In this case the current of hydrogen from the generator had been stopped and the pressure brought to 0·5 mm. On starting the discharge only the speck at *c* became incandescent. The pump was then started, and, after a few revolutions, the particle at B lit up, and a few moments later the particle at A. The pump was then stopped and the pressure again measured. It had fallen to 0·15 mm. Cooling a portion of the tube at *c* with a cotton pad wet with liquid air, immediately extinguished all of the glowing thoria particles in the tube above *c*, which is in perfect agreement with an observation by Langmuir, that atomic hydrogen from a glowing tungsten wire will pass for a considerable distance along a glass tube, but will not pass a point cooled with liquid air.

The distance which the atomic hydrogen travelled along the tube to the pump, as shown by the glowing thoria, is in good agreement with the earlier observation that re-combination takes place in about one-fifth of a second, for the velocity of the gas down the tube is of the order of a metre per second.

It would be interesting to determine the life of atomic hydrogen in the absence of any catalysing surface, for the wall of the tube, even under the best conditions, appears to have some action (as shown by our inability to get rid of a faint trace of secondary spectrum).

I intend to try the conditions in a large glass bulb placed in the pump circuit as close to the discharge tube as possible. If the life of the atomic hydrogen is longer under these conditions, a speck of thoria at the centre of the bulb should continue to glow for a measurable time after the discharge is stopped.

There appears to be strong evidence that the discharge tube is filled with practically pure atomic hydrogen, not only during the flow of the current, but also during the brief intervals of low potential between the half-cycles of the alternating current, when no current is flowing in the tube. This suggests the possibility of determining the optical properties (refraction, dispersion, etc.) of atomic hydrogen, by including the discharge tube in one path of an interference apparatus, illuminated by light which is passed through two apertures in a disc mounted on the shaft of a synchronous motor. The illumination of the fringe system will then occur only when the gas in the tube is non-luminous.

Summary.

It is shown that some metals, oxides and other substances are raised to incandescence when introduced into a stream of atomic hydrogen, the surface of the substance acting as a catalyser in bringing about the recombination of the atoms.

Atomic hydrogen, practically free from molecular hydrogen, can be drawn by a pump from the central portion of a long vacuum tube, excited by a high potential current. Fire-polished glass surfaces, such as the wall of a

glass tube, have a comparatively feeble catalysing power; fractured surfaces, however, cause the recombination of the atoms, and are strongly heated.

The action of water-vapour or oxygen in enhancing the Balmer spectrum, and suppressing the secondary spectrum of hydrogen, is probably due to its action on the wall of the tube, which, when dry, catalyses the atomic hydrogen as fast as it is formed by the current.

The peculiar spectroscopic phenomena observed with long hydrogen tubes, and described in an earlier paper, are explained.

Methods are suggested for determining the physical and optical properties of atomic hydrogen gas.

On the Scattering of β-Rays.

By H. A. Wilson, F.R.S., Rice Institute, Houston, Texas, U.S.A.

(Received May 23, 1922.)

The scattering of β-rays by thin sheets of matter has been the subject of a number of experimental and theoretical investigations in recent years. J. A. Crowther and B. F. Schonland* give a good summary of the principal results so far obtained, together with an account of a new series of measurements.

Crowther and Schonland regard their experiments as decisively in favour of the theory of "single scattering," which was shown by Rutherford to account for the scattering of α-rays by thin plates in a thoroughly satisfactory way. In the case of α-rays the fraction of the rays scattered through considerable angles is very small, and it is on this that the success of the theory depends. In the case of β-rays the fraction scattered is very much larger, and it is difficult to believe that "multiple scattering" can be neglected.

On Rutherford's theory of single scattering each ray is supposed to be deflected by a single collision with an atom. The number of collisions is clearly proportional to the thickness (t) of the sheet, so that the fraction scattered through more than a given angle is also proportional to the thickness and

$$I/I_0 = 1-kt, \tag{1}$$

where I denotes the number of rays deflected through less than a given angle, I_0 the number of rays incident on the plate, and k is a constant.

---

* 'Roy. Soc. Proc.,' A, vol. 100, p. 526 (February, 1922).

The constant k is found to be proportional to the square of the cotangent of the scattering angle (ϕ), so that if we take this angle very small, equation (1) makes I/I_0 large and negative, which is impossible. This equation therefore only applies to large angles of scattering. Small deflections are neglected on the single scattering theory.

So long as I and I_0 remain nearly equal and ϕ is not too small, equation (1) is satisfactory, but it is easy to see that it must be replaced by

$$I/I_0 = \epsilon^{-kt} \tag{2}$$

when I differs appreciably from I_0. Thus in one of Crowther and Schonland's experiments two gold foils scattered about 30 per cent. of the β-rays, and one foil scattered about 14 per cent. Now, if the first foil scatters 14 per cent. of the incident rays, then only 86 per cent. remain to be scattered by the second foil, so that we should expect the second foil to scatter 14 per cent. multiplied by 86/100, or 12 per cent. and not 14 per cent. Of course, the 14 per cent. scattered by the first foil passes through the second foil, and we should expect 14 per cent. of these, or 2 per cent., to be scattered a second time. This second scattering, however, is, roughly speaking, as likely to decrease the deflections as to increase them, and so can be neglected.

It is clear that when a considerable percentage of the rays is scattered then the exponential equation will represent the single scattering theory better than the linear equation. The two equations, of course, become identical for large values of ϕ when the fraction of the rays scattered is very small, as is the case with α-rays. Equation (2) also has the advantage that it makes I/I_0 equal to zero when $\phi = 0$, which is clearly correct.

When a β-ray passes through a thin sheet we should expect it to make a large number of collisions with atoms, so that we should expect scattering, due to the cumulative effect of many deflections. This idea is the basis of the "multiple scattering" theory which was put forward by J. J. Thomson in 1910. J. J. Thomson based his calculations on the theory that an atom consists of a solid sphere of positive electricity, inside which negative electrons move about, but the equation obtained is independent in form of any special theory as to atomic constitution. J. J. Thomson obtained the equation

$$I/I_0 = 1 - \epsilon^{-\phi^2/kt}, \tag{3}$$

where I/I_0 is the fraction of the β-rays remaining inside a scattering angle ϕ after passing through a plate of thickness t.

Crowther and Schonland considered that their results were decisively in favour of equation (1) as against equation (3), and conclude that the "single scattering" theory represents the scattering of β-rays when I/I_0 is greater than one-half.

According to what has been said, the equations which represent the two theories should be (2) and (3) and not (1) and (3), so I have compared the results given by Crowther and Schonland for gold with (2) and (3).

To do this the values of $t^{-1} \log I_0/I$ and $t \log I_0/(I_0 - I)$ were calculated.

(1) I/I_0. (Obs.)	(2) t.	(3) $t^{-1} \log I_0/I$.	(4) I/I_0. (Calc.)	(5) $t \log I_0/(I_0-I)$.	(6) I/I_0. (Calc.)
0·86	1	0·064	0·83	0·85	0·89
0·67	2	0·087	0·69	0·96	0·67
0·52	3	0·094	0·57	0·96	0·52
0·46	4	0·084	0·47	1·07	0·43
0·38	5	0·084	0·39	1·03	0·36

The first column of the above table contains the values of I/I_0 obtained from the curve given by Crowther and Schonland for $\phi = 0{\cdot}11$. The second column gives the number of gold foils through which the β-rays were passed. The numbers in the third column should be constant if equation (2) represents the results, and the numbers in the fifth column should be constant if equation (3) represents the results. It appears that either of the two equations can be used to represent the results about equally well.

The mean value of $t \log I_0/(I_0 - I)$ is 0·97, and if we put this value in the equation (3) of the multiple scattering theory and use it to calculate the values of I/I_0, we get the numbers given in the last column of the table. The mean value of $t^{-1} \log I_0/I$ is 0·0826, and using this with equation (2) to calculate I/I_0 we get the numbers given in the fourth column.

Both sets of calculated values agree fairly well with the observed values given in the first column, but the numbers calculated by the equation of the multiple scattering theory agree rather better with the observations than those calculated by the single scattering equation.

It seems quite clear that these results cannot be regarded as disproving the multiple scattering theory.

The following table contains a similar comparison of Crowther and Schonland's results with $\phi = 0{\cdot}145$ with equations (2) and (3):—

I/I_0. (Obs.)	t.	$t^{-1} \log I_0/I$.	$t \log I_0/(I_0-I)$.	I/I_0. (Calc.)
0·90	1	0·046	1·00	0·97
0·82	2	0·043	1·44	0·81
0·72	3	0·048	1·66	0·67
0·61	4	0·054	1·63	0·57
0·51	5	0·061	1·55	0·49

In this case the agreement between the numbers in the last column, calculated by equation (3), with the observed values, is not as good as in the first case; but equation (2) also does not represent the observations very well, as is clear from the way in which the values of $t^{-1} \log I_0/I$ increase with t.

The values of I/I_0 in this case agree nearly with equation (1), which does not represent either of the two theories, except when I/I_0 is very nearly equal to unity. It seems fair to conclude that these results do not enable us to decide between the two theories, but that they agree rather better with the multiple scattering theory than with the single scattering theory.

In view of this, it seemed to the writer worth while to work out the theory of multiple scattering of β-rays on the basis of Rutherford's nucleus theory of the atom, and to compare Crowther and Schonland's results with this theory as well as with the single scattering theory.

The theory of the motion of an electron near a positive nucleus has been worked out by C. G. Darwin, on the assumption that the total energy of the electron, when moving with a velocity v, is equal to $m(1-\beta)^{-\frac{1}{2}}$, where m is its energy when $v=0$ and is assumed constant, $\beta=v/c$, and c denotes the velocity of light. This formula gives the variation of the energy with v when the electron is far from other charged bodies, but when it is very near the positively charged nucleus, m is diminished and so does not remain constant.

Suppose an electron, consisting of a sphere of electricity of amount e and radius a, is at rest at a distance r from a positive nucleus, consisting of a sphere of electricity of amount E and radius b. The electrostatic energy of the system is then

$$\frac{1}{2}e\left(\frac{e}{a}+\frac{E}{r}\right)+\frac{1}{2}E\left(\frac{E}{b}+\frac{e}{r}\right).$$

In addition there is the internal energy of the charges which is $\frac{1}{6}\frac{e^2}{a}+\frac{1}{6}\frac{E^2}{b}$. The total energy is therefore

$$\frac{2}{3}\frac{e^2}{a}+\frac{2}{3}\frac{E^2}{b}+\frac{eE}{r}.$$

The energy $\frac{2}{3}\frac{e^2}{a}$ is located very close to the electron since its density varies inversely as the fourth power of the distance and the same is true of $\frac{2}{3}\frac{E^2}{b}$. The energy eE/r, however, is distributed in a symmetrical manner about the two charges, and its density very near either of them varies inversely as the square of the distance. This energy cannot therefore be regarded as moving with the electron like the energy $\frac{2}{3}\frac{e^2}{a}$.

When r is small most of the energy Ee/r is near the charges so that since its distribution is symmetrical it may be approximately true to suppose one-half of it to be associated with the electron and to move with it, and the other half to be associated with the nucleus.

Making this approximate assumption we have the energy of the electron, when it is at rest, equal to $\frac{2}{3}\frac{e^2}{a}+\frac{Ee}{2r}$, so that when the electron is moving with a velocity v its energy may be taken to be

$$\left(\frac{2}{3}\frac{e^2}{a}+\frac{Ee}{2r}\right)(1-\beta^2)^{-\frac{1}{2}}.$$

Suppose an electron starts towards a nucleus from a great distance with an initial velocity v_0 so that its initial energy is $m(1-\beta_0^2)^{-\frac{1}{2}}$ where $m=\frac{2}{3}\frac{e^2}{a}$ and $\beta_0 = v_0/c$.

Then assuming the total energy remains constant and that the nucleus is fixed we have

$$m(1-\beta_0^2)^{-\frac{1}{2}} = \left(m+\frac{Ee}{2r}\right)(1-\beta^2)^{-\frac{1}{2}}+\frac{Ee}{2r}.$$

The kinetic energy of the electron is equal to

$$\left(m+\frac{Ee}{2r}\right)\left((1-\beta^2)^{-\frac{1}{2}}-1\right) = m\left((1-\beta_0^2)^{-\frac{1}{2}}-1\right)-\frac{Ee}{r}.$$

Thus the kinetic energy on this theory has the same value as on the usual form of the theory.

The angular momentum of the electron about the nucleus is equal to

$$\frac{m+l}{c^2\sqrt{(1-\beta^2)}}\,r^2\dot{\theta} = p,$$

where $l = Ee/2r$, and θ is the angle r makes with a fixed radius in the plane of the orbit.

The energy equation gives

$$\sqrt{(1-\beta^2)} = \frac{m+l}{m'-l}$$

where $m' = m(1-\beta_0^2)^{-\frac{1}{2}}$ so that

$$p = \frac{1}{c^2}(m'-l)\,r^2\dot{\theta}.$$

On working out the equation of the orbit in the usual way we obtain

$$u = -\frac{Ee}{2p^2c^2}(m+m')(1+\epsilon\cos\theta)$$

where $u = r^{-1}$ and ϵ denotes the eccentricity. Thus, the orbit in the case considered will be a hyperbola, just as though the mass of the electron remained constant.

Thus, on the present theory, the complications of the usual theory do not appear even when the velocity approaches that of light.

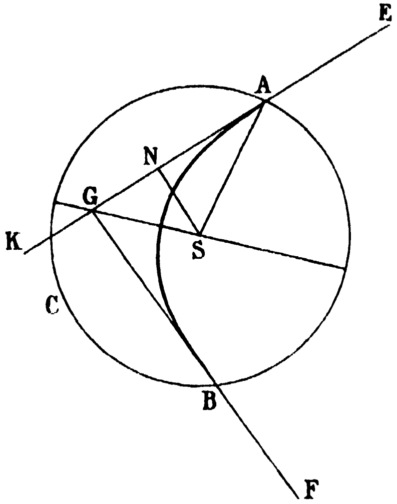
The kinetic energy of the electron has the usual value, but its total energy is equal to $m(1-\beta_0^2)^{-\frac{1}{2}}-\mathrm{E}e/2r$ or $(m+\mathrm{E}e/2r)(1-\beta^2)^{-\frac{1}{2}}$ instead of $m(1-\beta^2)^{-\frac{1}{2}}$.

According to J. J. Thomson's theory of multiple scattering the number of electrons remaining inside an angle ϕ is given by $\mathrm{I}/\mathrm{I}_0 = 1-\epsilon^{-\phi^2/kt}$ where kt is equal to the average value of the sum of the squares of all the deflections of a β-ray while passing through a plate of thickness t.

If we assume that each atom consists of a nucleus with a positive charge, equal to the charge on one electron multiplied by the atomic number N, surrounded by N electrons, and neglect the field due to the electrons, then the scattering on the multiple theory comes out very much greater than that observed. It is clear that the electric field of each nucleus cannot be supposed to extend to infinity, but must be regarded as extending only as far as the surrounding electrons. The electrons limit the field of the nucleus and so diminish the scattering.

To get an approximate theory of scattering I have assumed each atom to consist of a positive nucleus with a charge Ne surrounded by a hollow sphere of negative electricity of amount Ne and radius R. The field inside R is then that due to the nucleus, and outside R it is zero.

Actually the electrons are probably arranged in several layers of different radii. The value of R obtained by a comparison of the theory with the observed scattering will, therefore, be a sort of average value. Also if some of the electrons are very near* to the nucleus the β-rays may not pass inside them, so that the scattering may be the same as that due to an atom with a nuclear charge less than Ne.



In the figure let the nucleus be at S, and let the circle ABC represent the sphere of negative electricity of radius R. Also let EABF be the path of a β-ray. Let EA and FB produced meet at G, so that the angle FGK is the angle of deflection ϕ. Let SN be perpendicular to EK, and let $SN = h$.

Let the equation of the path inside the sphere be $Lr^{-1} = 1+\epsilon\cos\theta$, so that the equation of EA, which is the tangent to the conic at A, is $Lr^{-1} = \epsilon\cos\theta+\cos(\theta-\alpha-\phi/2)$, where α denotes the angle ASN.

At A, $r = R$ so that $LR^{-1} = \epsilon\cos(\alpha+\phi/2)+1$. If $r = \infty$ we get $\theta = \frac{\phi}{2}+\frac{\pi}{2}$ and

$$0 = \epsilon\cos\left(\frac{\phi}{2}+\frac{\pi}{2}\right)+\cos\left(\frac{\phi}{2}+\frac{\pi}{2}-\alpha-\frac{\phi}{2}\right) \quad \text{or} \quad \epsilon = \frac{\sin\alpha}{\sin\phi/2}.$$

Hence
$$LR^{-1} = 1+\frac{\sin\alpha}{\sin\phi/2}\cos(\alpha+\phi/2).$$

But $\cos\alpha = h/R$ so that

$$L = \frac{h^2}{R}+h\sqrt{\left(1-\frac{h^2}{R^2}\right)}\cot\frac{\phi}{2}. \qquad (4)$$

Now $L = -2p^2c^2/Ee(m+m')$, where m' denotes the total energy of the electron at $r = \infty$ in the hyperbolic orbit. The initial energy of the electron when the nucleus is surrounded by a negative sphere is equal to the energy in the orbit at $r = R$, and so is given by

$$m'' = m'-l'$$

where $l' = Ee/2R$.

Hence
$$L = \frac{-2p^2c^2}{Ee(m+m''+l')}$$

and
$$p = m''v'h/c^2,$$

where v' denotes the initial velocity of the β-ray, so that its initial momentum is $m''v'/c^2$.

Hence
$$L = \frac{-2m''^2v'^2h^2}{Ee(m+m''+l')c^2} = -\frac{2Th^2}{Ee}$$

where
$$T = \frac{m''^2-m^2}{m+m''+Ee/2R}.$$

Unless R is less than 10^{-11} cm., $Ee/2R$ is negligible, so that $T = m''-m$, which is the initial kinetic energy of the β-ray.

Let $k = -Ee/2T$ so that $L = h^2/k$.

With (4) this gives

$$h = \left\{\frac{1}{R^2}+\left(\frac{1}{k}-\frac{1}{R}\right)^2\tan^2\frac{\phi}{2}\right\}^{-\frac{1}{2}}. \qquad (5)$$

When a β-ray passes through a plate of thickness t, containing n atoms per unit volume, the number of collisions it makes for which h is between h and $h+dh$ is $2\pi nth\,dh$. If ϕ is the deflection corresponding to h, then the sum of all the values of ϕ^2 is given by

$$\Sigma\phi^2 = \int_0^R 2\pi nth\,\phi^2\,dh.$$

Substituting the value found for h and putting $\frac{1}{2}\phi = x$, we get

$$\Sigma\phi^2 = 8\pi nta^2R^2 \int_0^{\pi/2} \frac{x^2 \sin x \cos x\,dx}{(\cos^2 x + a^2 \sin^2 x)^2}$$

where $a = \frac{R}{h} - 1$.

Integrating by parts we get

$$\Sigma\phi^2 = \frac{8\pi nta^2R^2}{a^2-1}\left\{\int_0^{\pi/2} \frac{x\,dx}{\cos^2 x + a^2 \sin^2 x} - \frac{\pi^2}{8a^2}\right\}.$$

Also
$$\int_0^{\pi/2} \frac{x\,dx}{\cos^2 x + a^2 \sin^2 x} = \frac{\log a + 1}{a^2} + \frac{\log a + \frac{1}{8}}{3a^4} + \ldots\text{*}$$

When a is large compared with unity we have therefore

$$\Sigma\phi^2 = \frac{8\pi ntR^2}{a^2}\left\{\log a + 1 - \frac{\pi^2}{8}\right\}.$$

Putting in $a = \frac{R}{k}$ and $k = -\frac{Ee}{2T} = \frac{Ne^2}{2T}$ we get

$$\Sigma\phi^2 = 2\pi nt\left(\frac{Ne^2}{T}\right)^2\left\{\log R + \log\frac{2T}{Ne^2} + 1 - \frac{\pi^2}{8}\right\}.$$

Then according to the multiple theory of scattering

$$I/I_0 = 1 - \epsilon^{-\phi^2/\Sigma\phi^2}.$$

An examination of the integral in the expression for $\Sigma\phi^2$ shows that on this theory when a is large nearly all the scattering is due to the numerous small deflections. Large single deflections are so rare that they contribute little to the value of $\Sigma\phi^2$.

When t is very small the number of collisions will be small, so that $\Sigma\phi^2$ will not be equal to the calculated value which has been got by supposing that there are very many collisions corresponding to the values of h between h and $h+dh$. But the average value of $\Sigma\phi^2$ for a large number of β-rays will still be equal to the calculated value.

The value of $\Sigma\phi^2/t$ will now be calculated for β-rays passing through gold. The density of gold is 19·33, and its atomic weight 197·2, so that, since the mass of one atom of hydrogen is $1{\cdot}65\times10^{-24}$ grm., we have for the number of atoms of gold per cubic centimetre

---

* I am indebted to Prof. P. J. Daniell, of the Rice Institute, for the evaluation of this integral.

$$n=\frac{19{\cdot}33}{197{\cdot}2\times1{\cdot}65\times10^{-24}}=5{\cdot}94\times10^{22}.$$

The kinetic energy of the β-rays used by Crowther and Schonland was that due to a potential difference of $4{\cdot}59\times10^{5}$ volts, or $1{\cdot}53\times10^{3}$ electrostatic units.

For gold, $N=79$, which, with $e=4{\cdot}78\times10^{-10}$, gives $(Ne^2/2T)^2=1{\cdot}53\times10^{-22}$. Hence, if $\Sigma\phi^2/t=K$, then

$$K=228{\cdot}4\,\{24{\cdot}97+2{\cdot}30\log_{10}R\}. \tag{6}$$

The average value of K given by Crowther and Schonland's results used earlier in this paper can be easily calculated.

With $\phi=0{\cdot}11$, their results give 0·97 for the mean value of $t\log I_0/(I_0-I)$. This gives

$$K=\frac{(0{\cdot}11)^2}{8{\cdot}48\times10^{-6}\times0{\cdot}97\times2{\cdot}303}=636$$

where $8{\cdot}48\times10^{-6}$ is the thickness of each gold leaf. With $\phi=0{\cdot}145$ in the same way,

$$K=736.$$

The mean of these two values is 686. Crowther and Schonland also give the value of ϕ^2/t_m for these rays for gold, where t_m is the thickness which scatters half the rays. With I/I_0 equal to one half, we have

$$K=\frac{\phi^2}{\log 2\,.\,t_m},$$

Crowther and Schonland give $\phi^2/t_m=470$, which gives $K=680$, and so agrees nearly with the value calculated directly from their values of I/I_0.

If we assume $K=683$, we can calculate R by means of equation (6). The result is $R=2{\cdot}2\times10^{-10}$ cm. With this value of R, the theory of multiple scattering agrees with the scattering observed by Crowther and Schonland.

This value of R is probably too small, since the radius of an atom of gold is probably about 10^{-8} cm. It is probable, therefore, that an appreciable fraction of the electrons in the atom are inside most of the orbits of the β-rays. Let f denote the fraction of the electrons which are outside the orbits, so that fNe is the effective value of the nuclear charge. Then we have by (6)

$$K=228{\cdot}4f^2\,\{24{\cdot}97+2{\cdot}30\log_{10}R-2{\cdot}30\log_{10}f\}.$$

With $K=683$ this gives

$$\log_{10}R=-10{\cdot}84+\frac{1\,.\,3}{f^2}+\log_{10}f.$$

The following Table gives some values of f and the corresponding values of R:—

f.	R.
1	$2{\cdot}2 \times 10^{-10}$
0·7	$3{\cdot}3 \times 10^{-9}$
0·65	$1{\cdot}1 \times 10^{-8}$
0·60	$3{\cdot}6 \times 10^{-8}$
0·50	$1{\cdot}2 \times 10^{-6}$

Thus it appears that R varies very rapidly with f, and is equal to about 10^{-8} cm. when $f = 0{\cdot}65$.

When $f = 0{\cdot}65$ the effective nuclear charge is $51e$. If, then, we suppose that 51 of the electrons are outside the β-ray orbits at an average distance from the nucleus of 10^{-8} cm. and the remaining 28 inside the orbits, the theory accounts for the observed scattering. This seems to indicate that most of the electrons are near the surface of the atom, for the greater part of the scattering is due to collisions in which h is nearly equal to R, so that the deflection is small.

It seems clear that the theory of multiple scattering is adequate to account for the facts considered.

The single scattering theory will now be discussed. The number of collisions for which h is less than a given value made by a β-ray in passing through a plate of thickness t is equal to $\pi h^2 nt$, where n is the number of atoms per unit volume. In the case of single scattering through considerable angles the influence of the sphere of negative electricity becomes negligible, so that equation (6) reduces to

$$h = -\frac{Ee}{2T}\cot\frac{\phi}{2}$$

which is the same as the equation used by Rutherford for α-rays.

If I_0 denotes the intensity of the incident rays, then

$$I/I_0 = \epsilon^{-\pi nt(Ee/2T)^2\cot^2\phi/2} = \epsilon^{-k't}.$$

The mean value of $t^{-1}\log I_0/I$ got from Crowther and Schonland's results for gold with $\phi = 0{\cdot}11$ is 0·0826, which gives $k'\phi^2 = 271$. Substituting the values of n, E, e, and T, we get

$$k'\phi^2 = \frac{\pi n N^2 e^4}{T^2} = 114.$$

If we suppose that the effective value of the nuclear charge is fNe, then

$$f^2 = 271/114, \text{ so that } f = 1{\cdot}54.$$

Thus it appears that the observed scattering is considerably greater than that calculated on the single scattering theory.

The two theories, of course, are not mutually exclusive. The single scattering theory is clearly correct for very thin plates and large scattering angles, and the multiple scattering theory is clearly correct for thick plates and small scattering angles. It ought to be possible to work out a theory which would apply to plates of any thickness and to any scattering angles.

The multiple scattering theory in the case of rather thin plates does not take into account the large deflections, and the single scattering theory neglects the small deflections. It seems, therefore, that a more complete theory can be obtained by taking the product of the expressions for I/I_0 on the two theories, so that

$$I/I_0 = (1 - \epsilon^{\phi^2/kt})\,\epsilon^{-k't\cot^2\phi/2}$$

where
$$k = 2\pi n f^2 \left(\frac{Ne^2}{T}\right)^2 \left\{\log\frac{2TR}{fNe^2} + 1 - \frac{\pi^2}{8}\right\}$$

and
$$k' = \tfrac{1}{4}\pi n f^2 (Ne^2/T)^2.$$

This equation would evidently agree with Crowther and Schonland's experiments, but would require a slightly smaller value of the factor f than 0·65. It is clear that the single scattering theory is only applicable when the factor $1 - \epsilon^{-\phi^2/kt}$ does not differ appreciably from unity. When the multiple scattering factor is one half the single scattering factor is about 0·9, so that the multiple scattering is then much more important than the single scattering.

Taking $f = 0{\cdot}6$ and $R = 1{\cdot}1 \times 10^{-8}$, the following values of the two factors in the above expression for I/I_0 are obtained.

t.	ϵ^{-3400t}.	$1-\epsilon^{-0{\cdot}0000205/t}$	I/I_0.
1×10^{-6}	0·997	1·00	0·997
3	0·990	1·00	0·990
10	0·966	0·87	0·84
30	0·903	0·50	0·45
100	0·712	0·19	0·13

The first column gives the thickness of gold in centimetres. The second column gives values of ϵ^{-3400t} which is the single scattering factor for gold when $\phi = 0{\cdot}11$. The third column gives the multiple scattering factor for the same case, and the last column the product of the two factors. Thus, for thicknesses up to 3×10^{-6} cm. multiple scattering can be neglected, but with thicknesses above 10^{-5} cm., the multiple scattering is greater than the single scattering. The values of I/I_0 in the last column indicate an amount of scattering nearly equal to that observed by Crowther and Schonland for gold with $\phi = 0{\cdot}11$.

It appears that when both single and multiple scattering are taken into account, then if we assume $f = 0{\cdot}6$ and $R = 1{\cdot}1 \times 10^{-8}$ the theory agrees with the observations.

In an earlier paper[*] Crowther described a number of measurements of the scattering of β-rays, and obtained values of $\phi/\sqrt{(t_m)}$ nearly double those found for the heavier elements in the more recent work. In the earlier work larger scattering angles were used, and Crowther and Schonland state that they have recently found that with larger angles the scattering increases, and becomes equal to that previously observed by Crowther.

Since, on the theory proposed in this paper, the scattering for gold observed by Crowther and Schonland can be explained by taking $f = 0{\cdot}6$, it follows that the larger scattering observed by Crowther, and recently confirmed, will require f to be nearly unity.

We may suppose that, as the scattering angle is increased, the β-rays approach nearer to the nucleus, and so get inside more of the electrons, so that the effective nuclear charge increases.

With the single scattering theory employed by Crowther and Schonland, the scattering calculated for gold was nearly equal to that observed in the more recent experiments, and less than half that observed in the earlier experiments, although the effective nuclear charge was supposed to have its full value, so that they could not explain the increase of the scattering with ϕ as being due to an increase in the effective nuclear charge.

It appears that the theory proposed in the present paper is supported by Crowther and Schonland's experiments. If the usual form of the theory of the motion of an electron near a positive nucleus is used, instead of the modified theory suggested here, the calculated scattering is increased, so that the factor f has to be taken smaller in order to account for the facts.

The modified theory has the advantage of simplicity, and may be approximately true for electrons moving with high velocities when near the nucleus, but the success of the Bohr-Sommerfeld theory of spectra shows that the modified theory is not applicable to electrons moving comparatively slowly. For the purpose of explaining the scattering of β-rays the modified theory appears adequate, and I used it here because of its simplicity, since it is clear that the usual theory and the modified theory explain the facts equally well and merely require different values of the factor f.

Crowther and Schonland state that in the case of aluminium and copper their results agree slightly better with the assumption that the mass of the electron is constant than with the theory that it varies with its velocity as required by the theory of relativity. The results of the modified theory used here are the same, as regards scattering, as those of the theory of constant mass so that Crowther and Schonland's results appear to support the modified theory to some extent.

---

[*] 'Roy. Soc. Proc.,' A, vol. 84, p. 226 (1910).

A Study of Catalytic Actions at Solid Surfaces. Part VIII.—*The Action of Sodium Carbonate in Promoting the Hydrogenation of Phenol.*

By E. F. Armstrong, D.Sc., F.R.S., and T. P. Hilditch, D.Sc.

(Received June 7, 1922.)

Although the action of poisons on catalysts resulting in retardation or prevention of chemical action is well understood, much mystery still attaches to the so-called promoters of catalytic action. There are now a large number of well authenticated cases, both in scientific and patent literature, in which the presence of another metal or of some salt or other substance along with the catalyst facilitates action. No general explanation of promoter action has been advanced though attempts have been made to explain special cases.

Rideal and Taylor,* dealing with the ammonia synthesis, consider that the "promoter" may alter the relative adsorption of the reacting substances so that the optimum ratio for the desired interaction is more nearly approached. Pease and Taylor† have given a useful summary of the literature on catalyst promoters, with a view rather to future systematic investigation than to the establishment of a general theory at the present stage. Bancroft‡ has suggested that, instead of changing the ratio of adsorption, the promoter activates one of the interacting bodies chiefly, whilst the catalytic agent activates the other, the activation consisting apparently in some rearrangement of the chemical forces in the activated molecule. Palmer§ has recently suggested that the accelerating action of small amounts of magnesia or manganese oxide on the dehydrogenating capacity of copper is connected with selective absorption in the infra-red region of the spectrum, the whole process being due to radiation.

Before seeking any general theory to explain so many diversified phenomena it appeared to us desirable to investigate certain specific instances, the more especially as it seemed probable that definite and simple chemical explanations would be found differing from case to case. This expectation has been fully realised in the examples which are the subject of this and the following papers, in which it is shown that the activity of the promoter is due to causes which may be termed essentially chemical rather than physical. We would, however, at this stage of the enquiry specifically refrain from generalising on the subject.

---

* 'Catalysis in Theory and Practice,' pp. 29–31 (1919).

† 'J. Phys. Chem.,' vol. 24, p. 241 (1920).

‡ 'J. Ind. Eng. Chem.,' vol. 14, p. 331 (1922).

§ 'Roy. Soc. Proc.,' A, vol. 101, pp. 175–186 (1922).

We have made a prolonged study of the hydrogenation of liquid phenol in presence of nickel. This proceeds readily at 150–190° C., the products being cyclohexanol and cyclohexanone. The phenol must be very drastically purified before consistent results can be obtained, and even so, it is rarely found that two separate preparations of phenol are hydrogenated at exactly the same rate. In all cases, therefore, we have conducted a complete series of experiments on the same sample of purified phenol.

Few of the many workers on the hydrogenation of aromatic compounds seem to have studied the process kinetically, but Vavon and Detrie* state that in presence of platinum black a yield of 66 per cent. of cyclohexanol is produced, cyclohexanone being formed intermediately, and apparently an equilibrium is set up in which phenol is still present.

We find that, under the best conditions, little or no unchanged phenol is left, and that, by the time the phenol has all been transformed, the product consists of about 95 per cent. cyclohexanol and the remainder cyclohexanone.

Even with pure phenol, however, the action is notably accelerated by the presence of small quantities of mildly alkaline salts. Of those investigated, anhydrous sodium carbonate has proved the most suitable, whilst sodium borate acts similarly, but less energetically. The addition of caustic soda is detrimental, whilst anhydrous sodium phenate, sulphate or acetate produce a feeble acceleration by no means comparable with the effect of the carbonate. The presence of water retards the hydrogenation of liquid phenol, and the stimulation is only observed when the anhydrous salts are employed.

The following experiments illustrate the effect of varying proportions of sodium carbonate on the rate of hydrogenation of phenol. One hundred grams of pure phenol were taken for each experiment and hydrogenated at 180° C. for 1 hour under 80 lb. per square inch pressure of pure hydrogen, in presence of varying amounts of nickel and of sodium carbonate. The products were analysed and the percentage of reduction products isolated in each case is given in the following Table.

Thus sodium carbonate exerts an accelerating effect which is dependent on the amount present, the hydrogenation being retarded if an excess of the carbonate is present; moreover, the optimum amount of carbonate is not given by any definite concentration with respect to the phenol, for it varies in each of the cases.

---

* 'Compt. Rend.,' vol. 172, p. 1231 (1921).

Nickel present	1·0 grm.	0·50 grm.	0·20 grm.
Sodium carbonate—			
Nil	84·0	74·4	50·0
0·025 grm.	-	—	70·5
0·050 „	90·7	89·3	94·7
0·075 „	—	—	74·2
0·100 „	94·3	93·0	76·1
0·125 „	-	96·3	
0·150 „		94·4	
0·200 „	100·0	88·8	42·9
0·400 „	100·0	69·4	31·1
0·600 „	89·3		

On the contrary, it is directly related to the amount of nickel present, as the following analysis of the preceding data shows:—

Weight of phenol.	Weight of nickel.	Optimum weight of carbonate.	Percentage Na_2CO_3 on nickel.
grm.	grm.	grm.	
100	1·0	0·2–0·4	20–40
100	0·5	0·1–0·15	20–30
100	0·2	0·05	25

Thus sodium carbonate exerts its maximum influence when present to the extent of about 25 per cent. of the weight of nickel used in each case. It should be observed that this is the proportion of sodium carbonate with respect to the whole of the metallic nickel present; the whole of this is not necessarily concerned in actual catalysis, as we have shown in a previous paper,* but on the other hand, since the same preparation of catalyst was used throughout each series, the propoition of active nickel present will be proportional to the total weight used in each case.

The mechanism by which the sodium carbonate assists the process has not been discovered. It is possible that in some way the carbonate is associated with the catalyst, and protects it against the acidic influence of the phenol. However this may be, there is no measureable amount of carbonate associated with the nickel at the close of the experiment; the filtered residue of catalyst has been examined and found to be practically free from sodium carbonate, which must therefore be distributed in the liquid product of the action. Furthermore, sodium carbonate has been found to have a deleterious effect when added during the hydrogenation of organic acids, the stimulating action having only been observed, so far, amongst phenols (phenol, *o*-, *m*-, and *p*-cresol, β-naphthol, and thymol).

---

* Part VI of this series, 'Roy. Soc. Proc.,' A, vol. 99, p. 494 (1921).

Whilst the precise action of the carbonate is thus still obscure, we have clear evidence to show that the stimulation produced is not a true acceleration, but the restoration of the nickel to its normal function by the suppression of a retardation. This is shown by a comparison of the mean rates of absorption of hydrogen (litres per minute) over successive 10-litre periods, when 100 grm. of phenol were hydrogenated in presence of 0·5 grm. of nickel and varying proportions of sodium carbonate.

Sodium carbonate.	Nil.	0·050.	0·075.	0·100.	0·125.	0·150.	0·200.	0·400.
litres.								
0–10	0·980	1·493	1·449	1·613	1·493	1·667	1·429	1·695
10–20	1·031	1·316	1·538	1·493	1·613	1·667	1·587	1·852
20–30	0·800	1·053	1·370	1·471	1·515	1·724	1·613	1·538
30–40	0·602	1·042	1·266	1·493	1·538	1·538	1·408	1·111
40–50	0·532	0·971	1·124	1·282	1·587	1·587	1·087	0·807

The volume of hydrogen absorbed per increment of time declines steadily in the absence of sodium carbonate, but the decline becomes less marked until, at the optimum concentration (0·125–0·15 grm.), the action is represented by an almost linear absorption-time relation, of the type which we have previously shown to be characteristic for the hydrogenation of organic compounds in the liquid state in presence of nickel.

We have obtained further evidence of this transition from a curvilinear to linear absorption-time relation, by studying the hydrogenation of phenol at varying pressures of hydrogen in absence of sodium carbonate, and in presence of the optimum amount of this salt (0·15 grm. Na_2CO_3 for 0·5 grm. nickel).

Experiments were conducted with 100 grm. of the same sample of phenol throughout at 180° C., and at pressures of 20, 40, 60, and 80 lb. per square inch. Since, at the reaction temperature, the vapour pressure of phenol is one atmosphere, the effective, or partial, pressure of the hydrogen was approximately that indicated by the gauge, but this is only a rough approximation, since the vapour-pressure of the products (cyclohexanol b.p. 160° C., cyclohexanone b.p. 166° C.) is greater than one atmosphere at 180° C.

We give, in tabular form, the "constants" for a unimolecular action, calculated with respect to the amount of unchanged phenol present, together with the mean rates of absorption of hydrogen in litres per minute over the same intervals (successive 10-litre periods of absorption).

(i) Phenol Alone—Mean Rate of Absorption per Minute.

Pressure (lbs. per sq. in.).	40.	60.	80.
Hydrogen absorbed (litres.)			
10	0·704	0·962	1·235
20	0·662	0·901	1·205
30	0·541	0·813	1·111
40	0·413	0·658	1·042
50	0·400	0·476	1·020

(i) Phenol Alone—Constants for Unimolecular Action.

Pressure (lbs. per sq. in.).	40.	60.	80.
Hydrogen absorbed (litres).			
10	0·0043	0·0060	0·0076
20	0·0046	0·0062	0·0081
30	0·0046	0·0065	0·0087
40	0·0046	0·0067	0·0094
50	0·0047	0·0067	0·0105

(ii) Phenol in Presence of Sodium Carbonate—Mean Rate of Absorption per Minute.

Pressure (lbs. per sq. in.).	20.	40.	60.	80.
Hydrogen absorbed (litres).				
10	0·347	1·042	1·449	1·852
20	0·360	0·807	1·429	2·083
30	0·299	0·820	1·408	1·818
40	0·292	0·787	1·408	2·000
50	0·219	0·763	1·316	1·754
60	—	0·714	1·087	1·515
70	—	0·538	0·781	1·087

(ii) Phenol in Presence of Sodium Carbonate—Constants for Unimolecular Action.

Pressure (lbs. per sq. in.).	20.	40.	60.	80.
Hydrogen absorbed (litres).				
10	0·0021	0·0064	0·0090	0·0114
20	0·0024	0·0061	0·0096	0·0131
30	0·0024	0·0064	0·0105	0·0140
40	0·0027	0·0070	0·0116	0·0160
50	0·0028	0·0079	0·0132	0·0179
60	—	0·0093	0·0153	0·0209
70	—	0·0123	0·0200	0·0271

The figures for phenol alone at the lower pressures give perfectly sound constants for a "unimolecular" action with respect to the phenol present, a result which we previously obtained* in the case of other phenolic compounds (iso-eugenol and coumarin) at atmospheric pressure, and which is indicative of catalyst poisoning induced by the organic compound. In presence of sodium carbonate, the experimental data are no longer expressed by a logarithmic curve, and the action is more nearly represented by a linear relation.

The transformation of a "logarithmic" into a "linear" type of action by means of the addition of the "promoter" sodium carbonate is, when taken in conjunction with our previous work, strong evidence that the action of the carbonate is a protective one with respect to the catalyst, rendering the latter free to exercise its normal function. At the same time, and in agreement with this evidence, the relation which we have established between the concentrations of carbonate and catalyst (p. 23) indicates that the amount of protective agent must be in proportion to the amount of catalyst present in order to effect the maximum acceleration.

At high pressures in absence of sodium carbonate the unimolecular "constant" also increases steadily, a result which indicates that the acceleration in rate due to hydrogen concentration is not accompanied by any increase in the rate of poisoning of the catalyst by the phenol; in other words, the catalyst has not time to get thoroughly poisoned before it has finished its work of hydrogenation.

Furthermore, the fact that unimolecular constants are produced in the other cases indicates that the poisoning effect is slow enough to be measured, and that it is a function of the amount of phenol present; and, combined with the foregoing observation, may be supposed to suggest that the actual toxicity is caused by an *association of too stable a nature* between the nickel and phenol or some product of the latter (one simple possibility is that the retardation is due to formation of nickel phenate).

We may add that the data obtained for hydrogenation of phenol at various pressures indicate that the influence of pressure is very similar to that on the hydrogenation of other oxygenated compounds;† the closest resemblance is found in such cases as the hydrogenation of citral or of carvone. Thus phenol behaves like other oxygenated compounds containing a number of centres of unsaturation open to hydrogenation.

Whilst we have not been able to state in detail the action of sodium carbonate in the hydrogenation of phenol, the results are of interest, in the present state of vagueness on the subject of "promoters," because they establish for one particular case—

---

* 'Roy. Soc. Proc.,' A, vol. 98, pp. 27-40 (1920).

† 'Roy. Soc. Proc.,' A, vol. 100, p. 248 (1921).

(i) The connection between optimum concentration of promoter and amount of *catalyst*, not of organic compound;

(ii) The transformation of the action from one typical of the presence of catalyst poisons (in absence of carbonate) to one almost of the usual type prevailing in the absence of any catalyst poison (in presence of carbonate);

(iii) The "acceleration" of the action in presence of the "promoter" is in this case not a true acceleration, but the removal of a retardation; this leaves the catalyst itself free to behave more normally.

A Study of Catalytic Actions at Solid Surfaces. Part IX.—*The Action of Copper in Promoting the Activity of Nickel Catalyst.*

By E. F. Armstrong, D.Sc., F.R.S., and T. P. Hilditch, D.Sc.

(Received June 7, 1922.)

One of the most interesting cases of promoter action is that in which a mixture of two substances, each of which functions as a catalyst for the same action, shows greater activity than either constituent alone.

Thus in the synthesis of ammonia an iron-molybdenum catalyst is stated to afford greater yields of ammonia than is given by either of the metals alone.* Again, the production of formaldehyde from methyl alcohol and air was found by Hochstetter to be more effectively performed in presence of a silver-copper alloy than by either silver or copper alone.† Carbon monoxide is oxidised more efficiently in presence of certain mixtures of oxides, for example "hopcalite," than in presence of the constituent oxides (manganese and copper oxides).‡

A somewhat different case concerns the simultaneous action of two catalysts exerting different functions.

Camphor can be converted to camphane by catalytic reduction over nickel to borneol, dehydration of the latter to camphene over alumina, and hydrogenation of camphene to camphane. The hydrogenation proceeds readily at or below 200° C., but the dehydration in presence of alumina takes place at 350° C. or above. If, however, a mixed nickel-alumina catalyst be prepared, the whole sequence of changes may be carried out in one step, and at 200° C.*

---

* 'Badische Anilin und Soda Fabrik,' German Patent No. 5833 (1911).

† U.S.A. Patents 1100076, 1110289.

‡ Lamb, Bray and Frazer, 'J. Ind. Eng. Chem.,' vol. 12, p. 213 (1920).

We have observed the same phenomenon in the hydrogenation of phenol vapour over a nickel-alumina mixture at 180° C., when considerable proportions of cyclohexane are produced by dehydration and reduction of cyclohexanol; the latter requires, however, a temperature of 350° C. before dehydration to cyclohexane occurs in the presence of alumina alone.

To these instances may be added the subject of the present paper, namely, the behaviour of mixed copper-nickel catalysts in hydrogenation. Dewar and Liebmann† stated that mixed copper-nickel catalysts were of higher activity than nickel itself, and could be prepared by reduction at the very low temperature of 170–180° C.

We would emphasize here the importance of an adequate standard of activity; before committing ourselves to the statement that a mixed catalyst is more active than one of its constituents, it is necessary to consider under what conditions its activity is measured. Thus, in the case of nickel, Dewar and Liebmann's statement will undoubtedly hold if the comparison is made between a mixed copper-nickel preparation and pure finely divided nickel alone. But it is common knowledge that the full activity of nickel is not displayed unless it is adequately distributed on a support; we have indicated in a previous communication‡ some of the conditions determined by the surface-area of the catalyst. If the comparison is made between nickel thus distributed to display its maximum efficiency§ and mixed copper-nickel similarly mounted, we have never found the activity of the usual type of reduced nickel to be surpassed by the latter.

The interesting point is, that by employing supported copper-nickel compounds prepared in a suitable manner it is possible to obtain production of free nickel by exposure to hydrogen at 170–180° C., whereas similarly supported nickel compounds alone are not reduced by hydrogen below 300° C., and not rapidly below 350–400° C.

Detailed investigation of mixed copper-nickel catalysts has shown that in order to obtain the reduction of the nickel at this relatively low temperature the copper and nickel must, prior to reduction, be intimately and chemically associated with each other.

---

* Ipatieff, 'Ber.,' vol. 45, p. 3205 (1912).

† British Patents 12981/1913 and 15668/1914.

‡ Part VI of this series, 'Roy. Soc. Proc.,' A, vol. 99, pp. 490–495 (1921).

§ We may point out, however, that although distribution over a large surface and reduction of apparent volume undoubtedly have a preponderating influence, we have no definite knowledge at present as to whether the constituents of the more suitable supports also exert a specific activating ("promoting") effect on the nickel. In the absence of positive proof of the latter point, we prefer to compare catalysts mounted on a support which is known to give a good surface distribution of the metal.

In the first place, simple grinding together of nickel and copper compounds is of little use; to secure any effect the latter must be precipitated together before reduction.

This conclusion is based on some experiments in which various mixtures of copper and nickel carbonates were suspended in highly-refined cotton-seed oil and treated with hydrogen at 180° C. for 3 hours. The original oil absorbed 108·7 per cent. of its weight of iodine and the percentage absorption of iodine (iodine number) of the treated oils indicates the extent to which hydrogenation had proceeded.

The experiments were made both with the simple carbonates and the same carbonates precipitated in the customary manner on kieselguhr, 5 grm. of each preparation being used in conjunction with 150 grm. of the purified cotton-seed oil:—

	Iodine absorption of product after 3 hours at 180° C.	
	Carbonates alone.	Carbonates on kieselguhr.
Copper carbonate alone	108·0	99·0
Copper carbonate and nickel carbonate intimately mixed by grinding	100·1	97·2
Copper and nickel carbonates precipitated together from solution	2·5	3·1

It was found that whilst, of course, the copper was completely reduced to the metallic state in all cases, the nickel was only reduced in the latter case (co-precipitation), and then only to the extent of about 10 per cent. of the total nickel present. This was confirmed by examination of the products obtained by reduction of the copper-nickel carbonates in hydrogen alone at 170°–180° C. As a rule, only 7–8 per cent. of the nickel present was found to be in the metallic state, one or two exceptional cases showing up to 12 and 15 per cent.

Further experiments with co-precipitated copper and nickel carbonates on kieselguhr were marked by several failures to obtain more than a very little catalytic activity in the product. The cause of these variations was traced by means of a somewhat involved series of experiments, which we need not give in detail, to slight modifications in the precipitation of the mixed carbonates. The final conclusion reached was that, in order to obtain the best activity, a somewhat incomplete precipitation of the copper by alkaline carbonate was essential, for it was found that, as the filtrates from the preparation were more deeply coloured, so the activity increased. Parallel with this, it was observed that the colour of the prepared catalyst (in the form of precipitated carbonate) varied from brown, or grey in an inactive preparation, to pale blue, or greenish-blue, in the more active catalysts.

This may be illustrated by a few figures showing the melting-point of cotton-seed oil after hydrogenation for 3 hours at 180° C., under comparable conditions, with the same weight of prepared copper-nickel carbonate of varying appearance.

Copper-nickel carbonate on kieselguhr.		Melting-point of cotton-seed oil after 3 hours' treatment.
Colour of preparation.	Colour of alkaline filtrates.	
Brown	Colourless	Liquid
Drab	,,	,,
Greenish-grey	,,	50°
Pale greenish-blue	Pale blue	56°
Pale blue	Deep blue	61–62° (saturated)

The various basic copper carbonates which it was evident were in question were then studied in the light of Pickering's classical work on the complex carbonates of copper,* and it was found that a very close connection existed between the precipitates giving rise to active catalysts, and the class of compounds termed by Pickering α-cupricarbonates. The latter, formulated as $Na_2Cu(CO_3)_2$, are characterised by solubility in alkaline carbonate solutions to an intense blue solution, which contains electro-negative copper and oxidises glucose, whilst in caustic alkali solutions they yield dark violet-blue liquids, which, in addition to the above reactions, also dissolve cellulose, and, on standing deposit malachite or other basic carbonate. On the other hand, basic carbonates of copper alone are insoluble in both mild and caustic alkali if the molecular ratio of copper oxide to carbon dioxide is 3 : 1, or over, whilst if it is less than 3 : 1 they are insoluble in alkaline carbonate, but give the deep violet-blue solution with aqueous caustic soda.

The various preparations which we had made (some fifteen in all) were triturated with cold 10-per-cent. sodium carbonate and 10-per-cent. caustic potash solutions, allowed to settle, and the clear liquor decanted and examined, when it was found that an exact correlation existed between Pickering's criteria for α-cupricarbonates, and the preparations which yielded the more active catalysts on reduction. The data obtained for the five preparations quoted in the preceding Table will serve to illustrate this:—

---

* 'Chem. Soc. Trans.,' vol. 95, p. 1409 (1909).

Colour of preparation.	Melting-point of cotton-seed oil after 3 hours' treatment.	Aqueous extract of preparation with	
		10 per cent. sodium carbonate.	10 per cent. caustic potash.
Brown	Liquid	Colourless, no copper present	Pale yellow, trace of electronegative copper.
Drab	„	Pale yellow, no copper present	Electronegative copper.
Greenish-grey	50°	Slight pale green, trace of electronegative copper	Transitory violet-blue, oxidised glucose.
Pale greenish-blue	56°	Pale bluish-green, contained electronegative copper, oxidised glucose	Deep violet-blue, depositing deep-blue crystals on standing; oxidised glucose and united with cellulose.
Pale blue	61-62° (saturated)		

The parallelism between activity and the tests for cupricarbonates is exact, and it may legitimately be inferred that the power of reducing nickel carbonate at 180° C. is connected in some way with the presence of electronegative copper in the precipitated material.

The substance present in the washed and dried preparation is clearly not sodium cupricarbonate, which is unstable in dilute solution; it seems more probable that under controlled conditions of precipitation the latter salt, formed from the copper salt and sodium carbonate, has given rise to a nickel cupricarbonate, which is likewise not very stable and may easily be resolved into a mixture of nickel carbonate and basic copper carbonates.

Thus we are led to the view that to obtain any degree of catalytic activity the nickel and copper must not simply be intimately mixed by being precipitated together, but must also be contained in the same complex carbonate molecule.

When at least a portion of the nickel present is thus combined with copper, it is reduced at 180° C., and is ready for service as a catalyst of hydrogenation. We venture to suggest that the apparently low temperature of reduction may be simply explained if the above evidence of a nickel cupricarbonate formation is accepted: the reduction of copper carbonate or oxide is strongly exothermic, so much so that, for example, it is very difficult to maintain the temperature of a compact mass of copper oxide below 180° C. once reduction has set in. This heat is equally liberated, of course, as the copper in each cupricarbonate molecule is reduced, and although it is rapidly conducted away, it is obvious that at the moment of reduction the temperature of the molecule itself is momentarily raised to a considerable degree, and consequently the actual temperature of reduction of the nickel combined in the same molecule as the copper must be considerably above 180° C. We believe that the mechanism of the reduction is simply that the necessary local heat to provide a sufficiently high temperature for the reduction of the nickel is forthcoming from the reduction of the copper, providing that the latter is in sufficiently close proximity to the nickel—and our results indicate that nothing short of combination in the same molecule provides sufficiently close proximity.

The "promoter" action in this particular case is therefore open to a relatively simple explanation, especially as, when comparison is made between nickel suitably distributed on a support and a copper-nickel preparation on the same support, we have never been able to observe any activity in the latter case beyond that of the standard nickel itself.

Of many tests of the two types of catalyst which we have made under comparable conditions, we may set on record the following, in which in each case one part of total nickel was present per 1,000 of cotton-seed oil hydrogenated.

After one hour's treatment with the plain nickel catalyst, the melting-point of the cotton-seed oil was 52½° C.; with two of the most active copper-nickel catalysts which we have prepared, the melting-points of the oil after treatment were 49° and 46° C. In these experiments, the test was made by taking the copper-nickel catalyst after a preliminary three hours' reduction under cotton-seed oil and employing it in a fresh quantity of the oil; to render the test comparable, our standard nickel catalyst was treated with cotton-seed oil for a similar period under the same conditions before being tested on the oil as above described.

In thus accounting for the "promoting" action of copper-nickel catalyst we do not suggest that all cases of promoter action by the addition of small quantities of various substances to catalysts are to be similarly, or equally simply, explained.

As we have stated previously, however, we are of opinion that a thorough study of individual and specific cases in which catalysts are stimulated is the best prelude to a subsequent generalised theory.

On the Absorption of λ 5460·97 A. *by Luminous Mercury Vapour.*

By Prof. J. C. McLennan, F.R.S., D. S. Ainslie, M.A., and
Miss F. M. Cale, B.A., University of Toronto.

(Received June 17, 1922.)

[Plate 1.]

1. *Introduction.*

Numerous investigators who have studied the structure of the mercury green line λ 5460·97 A. have found that it consists of a wide central unresolved band, flanked on either side by three or more clearly resolved components that are usually designated as satellites. Janicki[*] has shown, however, and so has Nagaoka,[†] that the central portion of the line can be resolved into well defined components when a Lummer plate of high resolving power is used by itself, or is crossed with another of the same kind, or with a high grade *échelon* spectrograph. From such investigations it would appear to be definitely settled that the line is produced by twelve separable and distinct wave-lengths. These components, as measured by Nagaoka, are shown in their relative positions, and with approximately their relative intensities, in fig. 1. The values of their relative wave-lengths, as measured by Janicki and Nagaoka, are given in Table I.

In their investigations on the structure of the line λ 5460·97 A., Janicki,[‡]

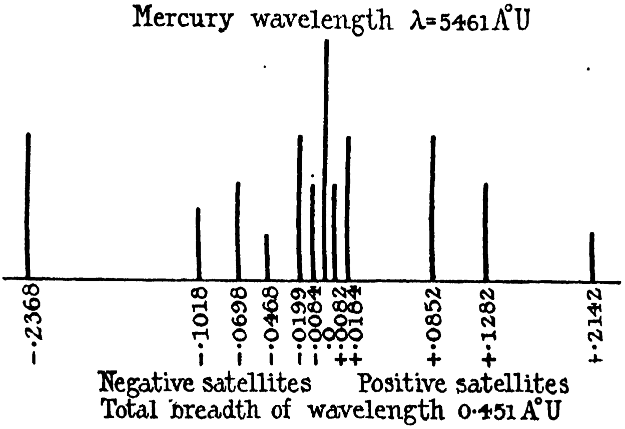


Fig. 1.

---

[*] Janicki, 'Ann. der Phys.,' vol. 33, p. 438 (1912).

[†] Nagaoka, 'Proc. Tokyo Math. Soc.,' 2nd Ser., vol. 8, No. 8, p. 229, October, 1915.

[‡] Janicki, 'Ann. der Phys.,' vol. 19, p. 35 (1906).

Table I.—Structure of Line λ 5460·97 A.

Component.	Δλ (Janicki).	Δλ (Nagaoka).	Δλ (Mean).
−6	−0·236 A.	−0·2368 A.	−0·2364 A.
−5	−0·102 A.	−0·1018 A.	−0·1019 A.
−4	−0·068 A.	−0·0698 A.	−0·0689 A.
−3	−0·048 A.	−0·0468 A.	−0·0474 A.
−2	−0·022 A.	−0·0199 A.	−0·021 A.
−1	−0·009 A.	−0·0084 A.	−0·0087 A.
M	0·000 A.	0·0000 A.	0·0000 A.
1	0·009 A.	0·0082 A.	0·0086 A.
2	0·018 A.	0·0184 A.	0·0182 A.
3	0·084 A.	0·0852 A.	0·0846 A.
4	0·1281 A.	0·1282 A.	0·1281 A.
5	0·2141 A.	0·2142 A.	0·2141 A.

Prince Galitzin and Wilip,* Stansfield† and one of the writers found that, when an Heraeus quartz mercury arc lamp was used as the source of light (and also when other forms of lamps were used as well), the structure of the line underwent a profound modification when the radiation was produced by the passage of very heavy electric currents in place of weak ones. Janicki observed, in place of the original line and its satellites, a peculiar system of five equidistant bands, the original components of the line being apparently lost in a continuous spectrum.

Prince Galitzin and Wilip suggested that the effect was probably due to a reversal of some of the constituents of the line or to some property of the resolving apparatus, and Stansfield rather favoured the view that the bands were due to secondary spectra produced by the *échelon* that he used. The view taken by one of us was that the phenomenon had the appearance of a reversal of the main component due to absorption, together with a widening and intensification of the satellites arising from an increase in the temperature and pressure of the mercury vapour in the lamp.

This view was in keeping with some observations made by Kuch and Retschinsky‡ on the illumination from two mercury lamps, one of which was placed behind the other so that the light from the first had to traverse the vapour in the second. They found that the radiation from the combination of lamps with this arrangement was less than the sum of the radiations from each separately. Similarly, from direct photometric measurements, it was shown by Pflüger§ and by Grebe|| that radiation of the wave-length λ 5460·97 A. had its intensity diminished by being passed through mercury vapour in a luminous condition.

---

* Prince Galitzin and Wilip, 'Bull. de l'Acad. Sc. de St. Pétersbourg,' 1907, p. 159.

† Stansfield, 'Phil. Mag.,' 1909.

‡ Kuch and Retschinsky, 'Ann. der Phys.,' vol. 22, p. 882 (1907).

§ Pflüger, 'Ann. der Phys.,' vol. 26, p. 789 (1908).

|| Grebe, 'Ann. der Phys.,' vol. 36, p. 834 (1911).

By the use of an interferometer of the Jamin type, Koch and Friedrich* showed that anomalous dispersion, too, could be obtained at λ 5460·97 A. with mercury vapour feebly luminous, but not with the vapour in the ordinary non-luminous condition.

Further, Starke and Herweg,† in reporting in 1913 on some experiments made by them preliminary to studying magnetic rotation and the inverse Zeeman effect with mercury vapour, state that they found, when the light from a strongly excited mercury lamp was passed through a second feebly excited one, they could, when a pair of nicols was used to suitably modify the intensity of the light emitted by the first lamp, obtain a reversal of the main component of the line λ 5460·97 A. They also state, without giving the wave-lengths involved, that they observed reversals of the satellites of the main line.

From this investigation it would appear that, in mercury vapour in the luminous condition, atoms of mercury exist in states suitable for the absorption of a number of those wave-lengths of light which in the aggregate produce the spectral line λ 5460·97 A.

In a paper published by Metcalfe and Venkatesachar,‡ in November of last year, there was described a series of interesting experiments on the absorption of the wave-length λ 5460·97 A. by mercury vapour rendered luminous by the passage of electric currents of varying intensities. In these experiments the observations were made with a Fabry and Perot interferometer, a low-power *échelon* spectrograph, and a concave grating having a radius of 10 feet and a ruling of 45,000 lines.

With the Fabry and Perot interferometer a reversal was obtained not only of the main component of the green line, but also apparently of the satellite λ + 0·09 A. (No. + 3, Table I). The authors cited also reported that, with the concave grating, the reversal of the green line on a continuous spectrum used as a background was obtained with striking clearness.

As Aston§ has shown that mercury consists of some six isotopes with atomic weights 197–200, 202, 204, it would follow that, in the spectrum of each of these isotopes, there should be a wave-length approximately equal to λ 5460·97 A. Moreover, if the green line in the spectrum of one of the isotopes of mercury consists of a main component and one or more satellites, one should expect to find a similar structure for the green line in the spectrum of each of the other isotopes. It is presumable also that in ordinary mercury vapour we have present atoms of each of the different isotopes of the element. A question naturally arises, then, as to the identity of the wave-lengths that constitute the green line as ordinarily observed. Do all of the wave-lengths shown in fig. 1 arise from atoms of one only of the isotopes of mercury, or do some of these wave-lengths originate in the atoms of different isotopes of mercury? Another question that arises is—should we expect the satellites of a main line to be absorbed to any appreciable extent when the main line itself is absorbed by luminous mercury vapour.

---

* Koch and Friedrich, 'Phys. Zeit.,' December 21, p. 1193 (1911).

† Starke and Herweg, 'Phys. Zeit.,' January 1, p. 1 (1913).

‡ Metcalfe and Venkatesachar, 'Roy. Soc. Proc.,' A, vol. 100, p. 149 (November, 1921).

§ Aston, 'Isotopes' (Edwin Arnold and Co.), 1921, p. 72.

If some of the components shown in fig. 1 originate in the atoms of different isotopes of mercury, then, on the basis of the results obtained by the investigators cited above, one should expect to find that on passing radiation of the wave-length λ 5460·97 A. through luminous mercury vapour reversals would be obtained, at least in the case of those wave-lengths which represent the main constituent in the line group of wave-lengths associated with each isotope.

For the purpose of investigating this matter further than has been done, a series of experiments was made by the writers at intervals as opportunity offered during the past few months, and as a result it has been found that only in the case of the main component, designated 0 in fig. 1, and of satellites Nos. +1 and −1, Table I, was complete absorption obtained by mercury vapour rendered luminous by the passage through it of electric currents. No reversals were obtained with any of the so-called satellites of the green line other than those mentioned. In particular, no appreciable absorption was observed in any of our experiments of satellite No. +3, $\Delta\lambda = +0·085$ A.

2. *Apparatus.*

In the experiments several different forms of mercury vacuum arc lamps were used, but one that was found specially useful is shown in fig. 2.

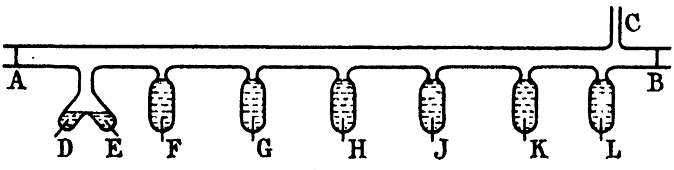


Fig. 2.

In operating this lamp, the current was drawn from the 110 D.C. mains. When the terminals D and E were used, the arc was easily started by giving the lamp a slight shake, and when other terminals were used, the lamp was generally started by heating the mercury in the terminal recesses with a bunsen burner and then passing a weak discharge from a small induction coil through the vapourised mercury.

In studying the constitution of the green line λ 5460·97 A., two Lummer plates and an *échelon* spectrograph of thirty plates were used, both already described in other communications.* In all experiments the radiation constituting the green line was isolated by means of one of the filters made by the Adam Hilger Company for the purpose.

3. *Emission Experiments.*

In the first set of experiments the constitution of the green line was investigated by analysing the radiation issuing laterally from a Cooper Hewitt mercury arc lamp, operated on currents of between 3 and 4 ampères.

The resolution obtained with the *échelon* spectrograph is shown in *b*, fig. 3, and with the Lummer plate in *b*, fig. 4, Plate 1. The constitution of the line, as revealed by the *échelon* and the Lummer plates, was a central unresolved band, flanked on one side by the satellites Nos. −6, −5, and −4, and on the other side by the satellites No. +3, +4.

From reproduction *b*, fig. 4, it will be seen that satellite No. −6 of one order in the Lummer plate diffraction pattern came between the unresolved band and satellite No. +3 of the next lower order. The correctness of the identification of this satellite was established by obtaining the fringe pattern with the *échelon* crossed with a Lummer plate. Such a fringe pattern is shown in fig. 5, Plate 1. In this pattern, satellite No. −6 is indicated by the letters *a, a, a, a.*

When the *échelon* was used crossed with a Lummer plate, it will be seen that satellite No. +5 (see *d, d, d, d,* fig. 5) was also brought into evidence. On several occasions in the course of our observations satellite No. −3 was seen, but it was always weak. It was faintly recorded between satellite No. −4, and the unresolved portion of the line on the negative from which *b* in fig. 4 was obtained. On account of its weak intensity, its distinctness was lost in the process of reproduction, and it only appears in the reproduction as a blurred strip.

As a result of our emission experiments we were able to confirm the existence of the satellites Nos. +3, +4, +5, and −3, −4. −5, −6. The main constituent and the satellites Nos. +1, +2 and Nos. −1, −2, we were not able to resolve directly with the optical equipment at our disposal. It will be seen later, however, that we were able to demonstrate quite definitely the existence of the satellites Nos. +2 and −2.

---

* McLennan, 'Roy. Soc. Proc.,' A, vol. 87, p. 269 (1912); McLennan and McLeod, 'Roy. Soc. Proc.,' A, vol. 90, p. 243 (1914).

4. *Absorption Experiments.*

In our first absorption experiments lamps similar to that shown in fig. 2 were used and observations were made by studying the radiation issuing from the ends of the lamp A and B. The lamp it may be stated was continuously exhausted during the experiments by a Gaede mercury pump. In preliminary experiments it was found that when a strong arc was established between D and E, D and F, D and G, F and G, or F and H, for example, the mercury vapour was distilled over from the arc towards B the colder portion of the tube. This distilled vapour was generally luminous and of a salmon or pinkish colour. Under these circumstances end-on observations at B showed that in the portion of the line $\lambda 5460{\cdot}97$ A. that we found unresolved in our emission experiment described above there always appeared a narrow dark reversal line. This reversal produced by the pinkish coloured luminous vapour was always distinct but it was not what could be called a strong reversal.

If, however, with a strong arc established between D and H, for example, we passed a discharge from a ¼ kilowatt Clapp Eastham 15,000 volt transformer operated on a current of about 1 ampère through the vapour between J and K or J and L it was found that reversal became much more marked. When the mercury in the recesses J and K or L, and the tube between these points as well, was strongly heated with the flame of a Bunsen burner, it was found possible with a certain application of heat, to obtain a vapour between J and K in such a condition of pressure or temperature, or probably both, that the appearance of the transformer discharge suddenly underwent a definite and very marked change. Under conditions of room temperature the discharge from the transformer was generally of a whitish appearance, but when the vapour reached the state mentioned, it suddenly took on a blue-greenish appearance. Under these circumstances the absorption was very marked, more extensive, and very clearly defined.

Photographs taken of the *échelon* and the Lummer plate patterns under these circumstances are reproduced in *a*, fig. 3, and in *a*, fig. 4, Plate 1. From these it will be seen that the central portion of the unresolved band in *b*, fig. 3, and *b*, fig. 4, was absorbed. This result was similar to what was observed by Metcalfe and Venkatesacher.

Various attempts were made by increasing the discharge from the transformer and by making alterations in the density of the vapour through which the discharge passed to widen out the absorption band, but it was found that while it could be obtained with widths extending all the way from a thin clearly defined line to the width shown in reproductions *a*, figs. 3 and 4, it apparently reached a definite width beyond which it could not be extended.

This suggested that what appeared in *a* of figs. 3 and 4 as two unabsorbed edges of the wide central band shown in *b*, figs. 3 and 4, was really the fringe pattern of the two satellites Nos. +2 and −2 found by Nagaoka and Janicki. To test this idea we carefully measured up our plates and it was found by taking satellite No. +4 as the base of our measurements that the unabsorbed edges gave for Δλ their reduced separation from the main component (as given by Nagaoka and Janicki) +0·019 A. and −0·020 A. As the averages of Nagaoka and Janicki's measurements of Δλ for satellites Nos. +2 and −2 were respectively +0·0182 and −0·021 we have concluded that what appeared as unabsorbed edges of the main line was in reality the two satellites Nos. +2 and −2.

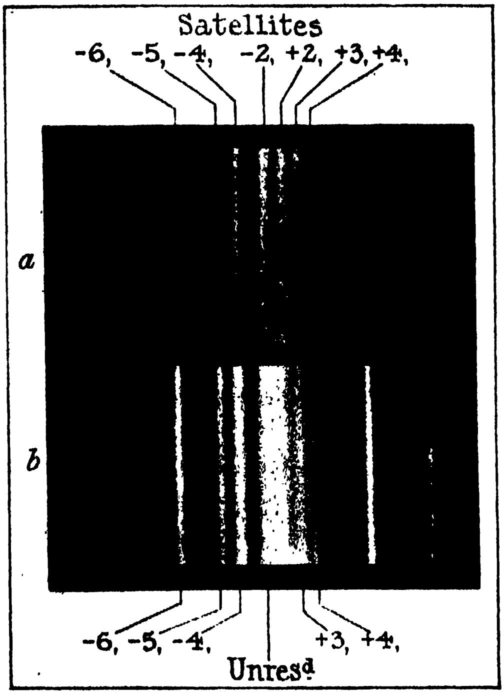


Fig. 3.

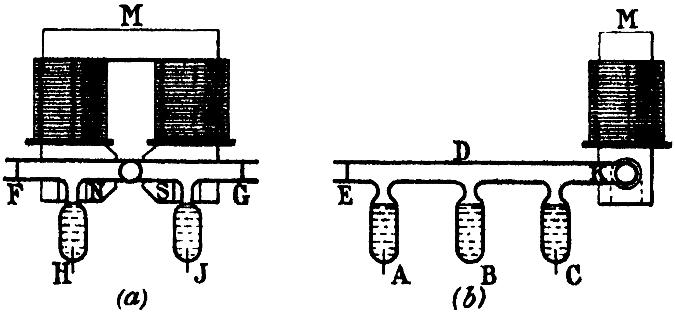
It would appear, therefore, from our experiments that under the maximum absorption conditions which obtained in our experiments the absorption of the main line and of the satellites Nos. +1 and −1 was complete.

Our experiments, therefore, showed that, with the conditions used by us, complete absorption was obtained of the main line and of the satellites Nos. +1 and −1, but no trace of reversals with any of the satellites Nos. +2 +3, +4, +5 or −2, −3, −4, −5 and −6. Plate 1, fig. 5, brings out the fact that when the main component and the satellites Nos. +1 and −1 were completely absorbed the satellite No. +5 was not absorbed. In this figure, *c* and *b*, *i.e.*, satellites Nos. +2 and −2, are shown quite distinctly with the components 0 and Nos. +1 and −1 absorbed. The figure also shows that under these conditions satellite No. 5, *i.e.*, *d*, *d*, *d*, *d*, did not show reversal, neither was it absorbed.

5. *Absorption of the Zeeman Components of the Green Line.*

From a number of investigations it is known that when a mercury arc is established in a weak magnetic field the radiation constituting the green line emitted at right angles to the field is magnetically resolvable into three components. In high magnetic fields, however, it can be broken up into nine equally spaced components, the spacing being equal to one quarter of the separation between the outer members of a normal Zeeman triplet. Of these nine components those constituting the central triplets are polarised in a plane perpendicular to the lines of force, while those constituting the two outer triplets are polarised in a plane parallel to the magnetic field.

In our experiments on the absorption of the magnetic components of the green line, the arrangement of the lamp used is shown in *a* and *b* of fig. 6 It consisted of the branches FG and EK. The one branch FG, shown in *a*, fig. 6, was inserted within the pole pieces of an electromagnet, so as to lie along the lines of force, and the other was attached centrally to it and at right angles, as shown in *b*, fig. 6. The arc was established between the terminals H and J, and the light which was studied was that which issued from the arc and passed through the branch KE. On issuing from KE the light was passed through a collimator, and into the Lummer plate in the usual way. On emerging from the Lummer plate, the light was passed through a Wollaston double-image prism and afterwards allowed to fall upon the lens of a camera. With this arrangement, the image in the focal plane of the camera consisted of two fringe patterns which overlapped over a portion of



(*a*) (*b*)

Fig. 6.

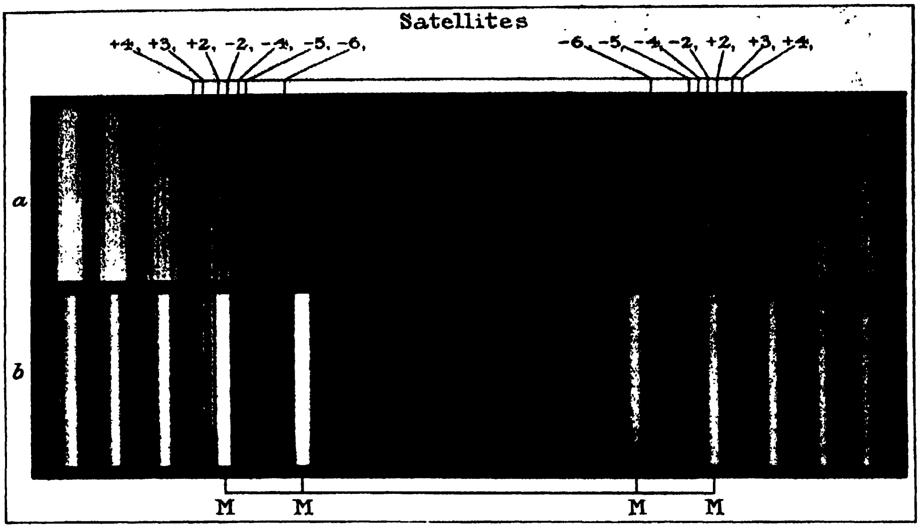


FIG. 4.

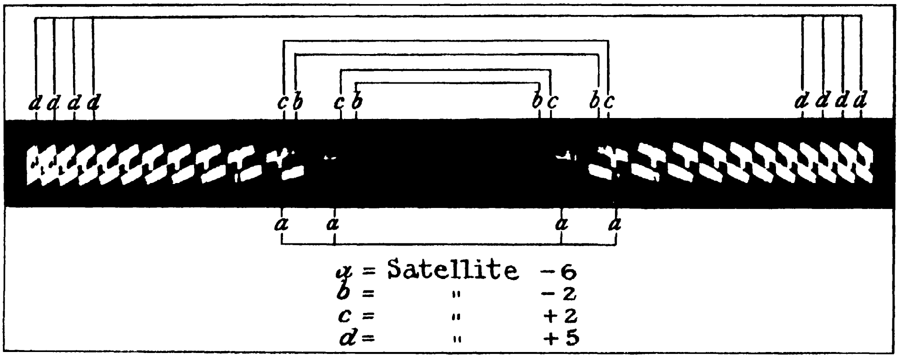


FIG. 5.

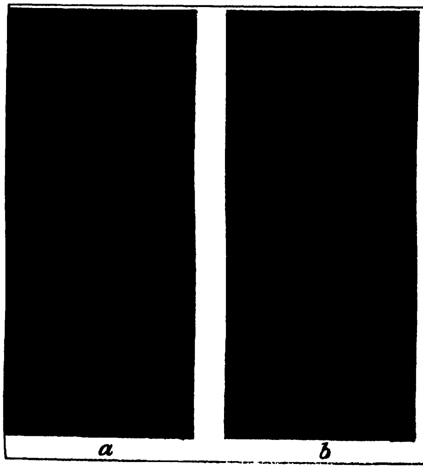


FIG. 7.

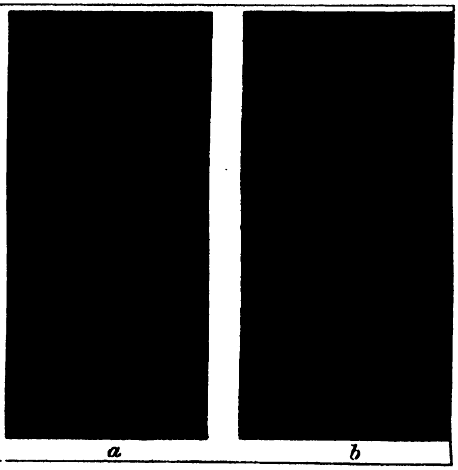


FIG. 8.

their length. Care was taken to give the Wollaston prism such an orientation that, with no magnetic field on, the fringes in the upper pattern were directly in line with those in the lower. The prism was also arranged so that the light forming the upper pattern was polarised in a plane parallel to the lines of force, and that forming the lower pattern polarised in a plane perpendicular to the field.

In the first experiment magnetic fields were used sufficiently high to resolve the green line into a Zeeman nonet. The method of coincidence was applied, and it consisted in so choosing the magnetic field that the outer triplet on the left-hand side of one order in the upper fringe pattern coincided with the outer triplet on the right-hand side of the next higher order of the same pattern. A reproduction of a photograph is shown in *a*, fig. 7, Plate 1. It will be seen from the reproduction how accurately coincidence can be obtained. The light forming the fringes in the upper pattern was, as already stated, all polarised in a plane parallel to the lines of force, while that forming the fringes in the lower pattern was all polarised in a plane perpendicular to the field. The undisplaced component of the nonet was therefore the central fringe in each of the triplets shown in the lower pattern. With this method of coincidence, it is easy to see that the separation $\Delta\lambda$ of each of the outer members of the triplets in the lower pattern relative to the undisplaced member was equal to $1/6\,\Delta\lambda_m$, where the value of $\Delta\lambda_m$ is given in the following table of data regarding the Lummer plate used to effect the resolution :—

Table II.

Data for Lummer plate.	Refractive Indices.	
	λ (A).	μ.
$d = 0·448$ cm.	6563·045	1·50746
$\lambda = 5461 \times 10^{-8}$ cm.	5896·155	1·50990
$\mu = 1·5121$	5890·186	
$d\mu/d\lambda = -530.$	4861·49	1·51500
$\Delta\lambda_m = 0·2984$ A.	4308·08	1·52025

Since $\Delta\lambda_m$ was equal to 0·2984 A., it follows that the separation of the outer components of the triplets in the lower pattern from the central component was 0·0497 A.

In the second experiment exactly the same method was followed as that just described, except that the branch of the lamp EK was kept warm with a bunsen burner, and the vapour in it was traversed by the discharge from the small transformer already described. The fringe pattern obtained in this case is that shown in *b*, fig. 7, Plate 1. There it will be seen the light forming the central undisplaced component was completely absorbed by the luminous vapour in the branch EK, while the light forming the other eight magnetic components of the line suffered no appreciable absorption.

In another set of experiments the magnetic field was so chosen that the two outer members of the outer triplet on the left-hand side of one order in the upper pattern coincided with the two outer members of the outer triplet on the right-hand side of the next higher order of the same pattern. With this arrangement, the fringe system in the upper pattern consisted of a set of quartets, while that in the lower pattern consisted of a set of triplets. Reproductions of the photographs of the fringe systems obtained without and with the discharge from the transformer passing through the vapour in the branch FG, are shown in *a* and *b* respectively in fig. 8, Plate 1. There, again, it will be seen that the obliteration of the central undisplaced component was complete. In this case, also, no evidence was obtained of any appreciable absorption by the luminous vapour in EK of the light forming the other eight magnetic components of the line. The separation $\Delta\lambda$ of each of the outer members of the triplets in the lower pattern relative to the undisplaced member was in this case $1/7\ \Delta\lambda_m$, *i.e.*, 0·0426 A.

In another set of experiments the magnetic field was so chosen that coincidence of the first class obtained. In this case the fringe system in the upper pattern consisted of a set of quintets, while that in the lower pattern again consisted of a set of triplets. Here again it was found that the central undisplaced component could be completely absorbed without the other eight magnetic components showing any appreciable absorption.

Finally, a magnetic field was used that was just sufficient to resolve the green line into a nonet. In this case there was no coinciding of the fringes, and both the upper and the lower patterns consisted of a set of very close triplets. Here also the central undisplaced component was completely absorbed without the other eight magnetic components showing any appreciable absorption. With this arrangement it was estimated that the separation of the outer components of the triplets in the lower pattern relative to the central undisplaced component was only slightly above 0·02 A.

With the evidence furnished by these experiments on the complete absorption of the central magnetic components of the line λ 5460·97 A. by luminous mercury vapour, it seems clear that we were justified in concluding that the fringes designated by +2 and −2 in *a* (fig. 3) and *a* (fig. 4, Plate 1) represented satellites Nos. +2 and −2, and not the unabsorbed edges of a broad and central main component of the line.

In this connection it may be pointed out that it follows from this conclusion that absorption by luminous mercury vapour of the light constituting the green line affords a means of easily and clearly demonstrating the existence of the satellites of the green line with displacements $\Delta\lambda = +0{\cdot}0182$ A. and $\Delta\lambda = -0{\cdot}021$ A.

It is of interest to note in passing that, in the absorption of the central magnetic component of the green line by the luminous vapour in FG, we have an example of plane polarised light being absorbed by atoms in a condition where they are capable of emitting only unpolarised light.

6. *Isotopes of Mercury.*

From what has gone before, it follows that, when light constituting the ordinary green line of mercury is passed through moderately luminous mercury vapour, the components of the line designated by zero and the two components separated from it by $\Delta\lambda = -0{\cdot}087$ A. and $+0{\cdot}0086$ A. can be completely absorbed.

In seeking for an explanation of this exceptional absorption by luminous vapour of these components, and the lack of any marked absorption in the case of the other components, one is reminded of the fact that mercury, as we ordinarily use it, consists very probably of five, and possibly of six, isotopes, with atomic weights 197–200, 202, 204.

It may be that the three components referred to originate in different isotopes of mercury. If this should be so, one would naturally associate the component zero with isotope 200 and the components $\Delta\lambda = -0{\cdot}0087$ A. and $\Delta\lambda = +0{\cdot}0086$ A. with the isotopes 202 and 198 respectively.

According to the Bohr theory, we have for two isotopes of an element the frequency relation

$$\frac{\nu_1}{\nu_2} = \frac{M_1(M_2+m)}{M_2(M_1+m)}, \quad i.e. \quad \frac{\lambda_2}{\lambda_1} = \frac{M_1(M_2+m)}{M_2(M_1+m)},$$

where M_1 and M_2 are the masses of the atoms of the respective isotopes and m is the mass of an electron.

i.e.

$$\frac{\Delta\lambda}{\lambda_1} = \frac{m(M_1-M_2)}{M_2(M_1+m)}.$$

Applying this to the case of isotopes 200 and 202 we have

$$\frac{\Delta\lambda}{\lambda_1} = \frac{-0{\cdot}0005\times 2}{202(200+0{\cdot}0005)} = \frac{-0{\cdot}001}{202\times 200{\cdot}0005} = \frac{1}{40400101},$$

or

$$\Delta\lambda = -\frac{1}{40400101}\times 5460{\cdot}97\text{ A.} = 0{\cdot}000135\text{ A.}$$

In a similar way the separations $\Delta\lambda$ corresponding to all the other isotopes can be calculated. The results are given in Table III:—

Table III.

Isotope.	Δλ.	Δλ × 80.
204	−0·000268 × 10^{-8} cm.	−0·02144 × 10^{-8} cm.
202	−0·000135	−0·01080
200	0·00000	0·00000
199	+0·0000685	+0·00548
198	+0·000138	+0·01104
197	+0·000279	+0·02032

From the above results it is seen that the maximum separation of the wave-lengths at or near 5460·97 A. arising from all the isotopes of mercury is, on the Bohr theory, approximately 0·000547 × 10^{-8} cm., *i.e.*, it is about 1/15 of the separation of either of the satellites Nos. +1 and −1 from the zero central component.

On the Bohr theory, then, none of the eleven satellites of the green line can be considered as originating in any of the isotopes of mercury with atomic weights 197, 198, 199, 202, and 204.

In some experiments, however, that we have recently made with lithium we have found that the red line of lithium consists of two doublets with a separation between three and four times that calculated on the basis of Bohr's theory for isotopes of atomic weight 6 and 7. From this result and from the results obtained by Merton* and by Aronberg† in studying the spectral displacements for isotopes of lead, we have been led to put forward the view that the observable spectral displacements for isotopes should be given by the atomic number times their displacements calculated on the Bohr theory. To test this view, the displacements calculated for the isotopes of mercury on the Bohr theory have been multiplied by 80, the atomic number of mercury, and the results are given in column 3 of Table III.

Table IV.

Isotope.	Isotope displacement × atomic number.	Satellite.	Satellite separation.
204	−0·02144 × 10^{-8} cm.	No. (−2)	−0·210 × 10^{-8} cm.
202	−0·0108	No. (−1)	−0·0087
200	0	M	0
199	0·00548	—	—
198	0·01104	No. (+1)	0·0086
197	0·02032	No. (+2)	0·019

These results are also given in Table IV, together with the separations of satellites Nos. ±1 and ±2, as found by Janicki and Nagaoka.

From this Table it will be seen that the calculated displacements for isotopes 204, 202, 198, and 197 are in approximate agreement with the displacements corresponding to satellites Nos. −2, −1, and Nos. +1, +2. No satellite of the green line, however, has been identified as yet with the displacement calculated for an isotope of atomic weight 199, and this is rather interesting, for while Aston has found that there are isotopes of mercury with atomic weights between 197 and 200, he has not as yet been able to assert definitely that one exists with atomic weight 199.

---

* 'Merton, 'Roy. Soc. Proc.,' A, vol. 96, p. 388 (1920).

† Aronberg, 'Proc. Nat. Acad. Sc.,' vol. 3, p. 710 (1917); and 'Astrophys. Journ., vol. 47, p. 96 (1918).

Too great emphasis, however, should not be placed upon the view we have put forward tentatively that the spectral displacements for different isotopes can be determined by multiplying by the atomic number the displacements calculated on the Bohr theory.

For even if this view should turn out to be correct, and it might be used to explain the absorption by luminous mercury vapour of the main component and of the components Nos. +1 and −1 of the green line, it would leave unexplained the non-absorption by luminous mercury vapour of the components of the green line Nos. +2 and −2.

7. *Summary of Results.*

1. It has been shown that when the radiation constituting the green line of mercury is passed through moderately luminous mercury vapour, the main component and the components No. +1, $\Delta\lambda = 0·0086$ A., and No. −1, $\Delta\lambda = -0·0087$ A., can be strongly absorbed.

2. No marked absorption by luminous mercury vapour was observed in the case of the other nine components of the green line.

3. Of the nine members constituting the magnetically resolved green line, it was found that the central undisplaced member was the only one that could be markedly absorbed by luminous vapour.

4. It has been shown that absorption by luminous mercury vapour of the light constituting the green line in the spectrum of mercury, affords a means of clearly and easily demonstrating the existence of the components of the line with separations $\Delta\lambda = +0·0182$ A., and $\Delta\lambda = -0·021$ A., *i.e.* satellites Nos. +2 and −2.

5. Some considerations have been presented in support of the view that the components of the green line of mercury, for which $\Delta\lambda = +0·0182$ A., 0·0086 A., −0·0087 A., −0·021 A., and zero, may originate in atoms of the element having respectively the weights 197, 198, 202, 204, and 200.

The Effect of Temperature on Gravitative Attraction.

By P. E. Shaw, B.A., D.Sc., and N. Davy, B.Sc., A.R.C.S.

(Communicated by Prof. E. H. Barton, F.R.S. Received January 9,—Revised July 24, 1922.)

Previous work by one of us* with a torsion balance of the Boys-Cavendish type had appeared to indicate that the so-called Newtonian Constant of Gravitation increases slightly as the attracting bodies are heated. The attempt has now been made to carry the work to a higher order of accuracy with a view to confirming or disproving that result. The clearing up of this point has become of greater importance on account of the recent development in the theories of gravitation.

The same delicate torsion balance was used as in the former experiments. This consisted of a fine quartz fibre carrying a light beam, from which two silver balls were suspended by fibres of unequal length. It was contained in a vacuum tube, and the observations were made by means of a beam of light proceeding from an illuminated scale through a window in the vacuum tube, and reflected by a small plane mirror attached to the torsion fibre. Two lead spheres, each of 100 lbs. weight, were placed near the silver spheres, so that by their attraction first to one side and then to the other, a torsional deflection might be obtained from which to calculate G. These deflections were not determined directly on account of the length of time required for the torsion system to come to rest, but were calculated from observations of successive turning points of the image of the scale.

The most important improvement in the new arrangement was in regard to the methods of suspension, both of the vacuum tube and of the large attracting spheres. Previously, trouble had arisen from disturbing tremors due chiefly to distant railway trains, which made accurate observations of turning points difficult to secure. The aim was to keep out tremors by arranging a suspended system, not too tightly gripped, yet held firmly enough to prevent slight displacements of component parts of the apparatus which would render the results worthless. For this purpose the vacuum tube was suspended from a cranked iron bar, in cantilever fashion, the bar itself being capable of turning about a horizontal axis, but held by steel springs and lead weights, so as to keep the tube vertical. This arrangement formed a "floating system," any lateral displacements in which were prevented by thick pads of indiarubber. Moreover, the floating system was itself suspended from strong angle irons heavily loaded, and resting on steel balls embedded in rubber placed in recesses in the walls of the building. With these dispositions it was found possible, under good working conditions, to estimate the turning points of the moving image of the scale accurately to 0·1 mm. A further improvement was obtained by increasing the distance from mirror to scale to 6 metres. Around the suspended system were wound coils of tubing through which a steady stream of cold water was passed, and this, together with a packing of cotton-wool, maintained the vacuum tube and its contents at a steady temperature.

---

* Shaw, 'Phil. Trans.,' A., vol. 216, 1916.

A temperature effect was sought in the differences of the deflections of the balance obtained when the large masses, used first cold, then hot, and then cold again, were moved from an A position to a B position. To give the results of one complete experiment (the whole number of which was eleven)—In millimetres of scale, the mean of seven "cold" deflections was 207·06 (greatest 207·20, least 206·95), and of twelve "hot" deflections 206·98 (greatest 207·15, least 206·85). Thus the mean difference in range was −0·08 mm., and as this is within the limits of experimental error, the result deduced is that, assuming G does vary according to a law $G = G_0(1+\alpha t)$, then α is at any rate numerically less than $1{\cdot}6 \times 10^{-6}$. Thus, as far as these experiments go, *i.e.*, for a temperature range 0° to 250° C., G remains constant.

This result is in opposition to that of the previous work, and we are of the opinion that slight displacements of the heavy masses occurring during the experiments introduced systematic errors in the former work. In the present experiments we think we have secured the absence of these by firmer support.

We wish to record our obligation to the Royal Society for a grant in aid of this research.

The Loss of Energy of an α-Ray Beam in its Passage through Matter. Part I.—*Passage through Air and CO_2.*

By P. L. Kapitza, Lecturer in Physics in the Polytechnic Institute, Petrograd.

(Communicated by Prof. Sir E. Rutherford, F.R.S.—Received June 21, 1922.)

Introduction.

An α-particle, emitted from a radio-active body, will, during its passage through matter, gradually lose kinetic energy. The experiments of Rutherford and Robinson* have shown that this kinetic energy is converted into an equivalent amount of heat generated in the matter.

These experiments confirm the law of the conservation of energy, but give no indication of the mechanism whereby the energy is given to the surrounding space. This mechanism has been examined theoretically by Thomson,† Bohr,‡ and Darwin,§ who proved that the energy of the α-ray, when passing near an atom, was first given to the electrons, which, in turn, yielded their energy to the atom. The manner in which the α-particle yields its energy to the electrons depends, on the average, on the speed of the α-particle, and on the number and arrangement of the electrons within the atom.

The dependence of the loss of energy of the α-particle on its speed and the structure of the atom has been examined by two methods. The first is that in which the α-particle is projected through different thicknesses of material, its emergent velocity being determined by the deflection produced in a magnetic field. This method has been employed by Rutherford,|| Geiger,¶ and Marsden and Taylor.** This is, *a priori*, the most exact method of attack, but in the experiments there is a weak point, in that the deflected beam must be observed, photographically, or by means of scintillations, and towards the end of the range, when the energy is small, these methods of detection lose their sensitiveness.

It is a surprising fact that, so far, no α-particle has been detected by these methods with a velocity less than 0·4 of the α-particle from radium C. Consequently, it is not known how the last 16 per cent. of the energy of

---

* E. Rutherford and H. Robinson, 'Phil. Mag.,' vol. 25, p. 312 (1913).

† J. J. Thomson, 'Conduction of Electricity through Gases,' pp. 370–382.

‡ N. Bohr, 'Phil. Mag.,' vol. 25, p. 10 (1913).

§ C. G. Darwin, 'Phil. Mag.,' vol. 23, p. 907 (1912).

|| E. Rutherford, 'Phil. Mag.,' vol. 12, p. 138 (1906).

¶ H. Geiger, 'Roy. Soc. Proc.,' A, vol. 83, p. 505 (1910).

** E. Marsden and T. S. Taylor, 'Roy. Soc. Proc.,' A, vol. 88, p. 443 (1913).

the α-particle is expended, in spite of the fact that this energy is quite considerable.

The second method consists in the study of the amount of the ionisation produced by the α-particle in different parts of its range. If we assume that all the energy of the α-particle is lost in the production of ions, and that the production of each ion requires the same amount of energy, then we can, from the ionisation curve, obtain information about the loss of energy of the α-particle in its passage through a gas. The weak points of this method are (1) that the assumption has no definite theoretical or experimental foundation, and (2) that the method cannot be applied to matter in the solid state.

In the present experiments, which were carried out at the suggestion of Sir Ernest Rutherford, an attempt is made to determine the energy of the α-particle by measuring the heating effect which it can produce after having travelled a definite portion of its range. Any method which is to be used for this purpose must possess the following qualities:—First, since it is necessary to use a narrow beam of α-particles, the energy of a small fraction (about 1000th) of the total number of α-particles emitted from the source has to be measured. If the measuring instrument has a sensitiveness of 10^{-9} calories per second, it would be possible to detect an energy of 0·3 per cent. of the initial energy by determining the range to an accuracy of 1 per cent. Second, if a source of RaC is used, which decays rapidly, the measuring instrument must, in order to obtain a considerable number of readings, have a small thermal capacity, and must also be able to reach a state of equilibrium rapidly. A Boys' radio-micrometer could be used for this purpose, provided it were made sufficiently sensitive and the construction modified to suit the above requirements.

General Theory of the Apparatus.

The directions of modification of the apparatus may be seen from the following considerations. The principal part of the radio-micrometer (fig. 1) consists of a loop of wire of high conductivity, L, terminated at one end by a thermocouple, JT, and supported by means of a quartz fibre, F, in a magnetic field, H. If one end, J, of the thermo-junction is heated, a current circulates in the loop, and a deflection is produced. Having determined the resistance, r, of the thermocouple and the loop (that of the latter being small), the area, S, of the loop, and the torsion-couple per unit twist, μ, of the fibre, an increase, ΔT, in the temperature of one end of the thermocouple produces a deflection, θ, given by

$$\theta = B\,\Delta T\,\frac{SH}{r\mu}, \tag{i}$$

where B is a constant depending on the nature of the thermocouple. At first sight, it appears that great sensitiveness could be produced by having S and H large, and μ small. But it is necessary to bear in mind the second condition, which demands that the time for the deflection to reach its maximum value should be small. Taking into account the time, if I is the moment of inertia of the suspended system, we have the following equation of motion—

$$\mathrm{I}\frac{d^2\theta}{dt^2}+\mathrm{K}\frac{d\theta}{dt}+\mu\theta=f(t), \qquad \text{(ii)}$$

in which K is the damping factor. This damping is a consequence of the eddy currents induced in the loop during its motion in the magnetic field. It is easy to show that K is given by

$$\mathrm{K}=\frac{\mathrm{H}^2\mathrm{S}^2}{r}. \qquad \text{(iii)}$$

If the roots of the characteristic equation of (ii) are imaginary, damped oscillations will be executed by the system. If the roots are real the maximum displacement is reached in a continuous manner. For the so-called critical point, when the roots are equal, the system attains its maximum displacement in the shortest time. In practice it is not necessary to work exactly at the critical point, but for the purpose of mathematical investigation we shall consider this particular condition. Then

$$\mathrm{K}^2=4\mathrm{I}\mu, \qquad \text{(iv)}$$

and the period of oscillation of the system, τ, is given by

$$\tau=2\pi\sqrt{(\mathrm{I}/\mu)}. \qquad \text{(v)}$$

Introducing (iii), (iv), and (v) into equation (i), we get

$$\theta_{\text{max.}}=\Delta\mathrm{TB}\sqrt{(\tau^3/4\pi^3\mathrm{I}r)}.$$

This solution shows that for a given τ, $\theta_{\text{max.}}$ is independent of S, and increases as I and r decrease. These considerations lead us to conclusions very different from those indicated at first sight by equation (i). It is necessary to make the loop small so that the moment of inertia may be small and also that the wire should not be too thin so that r may be kept small.

Further, by making the radio-micrometer small, it is easier to shield it from stray external temperature effects, while the time to reach the thermal equilibrium value is small. These two considerations were of the highest importance, for in the experiments it was found that it was much more difficult to shield from the stray effects than to increase the sensitiveness to any desired amount.

Description of the Apparatus.

(i) *Suspension System* (figs. 1 and 2).—The suspension system is placed between pole pieces, N, S, of circular cross-section and separated by a distance of 3·4 mm. It is supported by a short quartz fibre, of thickness 2·0 μ to 2·5 μ. The magnetic field is produced by a double horseshoe magnet, M, M, taken from a Bosch magneto. As pointed out by Boys in his earlier papers, the greatest difficulty in the construction of the suspension system arises from the fact that, owing to the magnetism of the loop, the motion cannot be controlled by the use of thin quartz fibres. In a recent paper, Witt* has given a very elegant method for the production of non-magnetic suspension systems. This method was applied in the following manner: A copper wire, diameter 0·14 to 0·23 mm., was taken and bent into a loop, L, as shown in fig. 1. After carefully cleaning with acid and washing with distilled water, it was attached to a quartz fibre which supported a mirror. It was then placed between the sharp-pointed pole-pieces of an electromagnet, between which a magnetic field with a large gradient could be obtained. When the magnet was excited, a very large deflection of the spot of light from the mirror occurred. This deflection was due to the paramagnetism of the specimen, owing probably to the presence of small amounts of iron generally found in commercial copper. Pure copper is diamagnetic. The loop was now placed in an electrolytic bath and pure copper deposited on it. Starting with a current of 1·5 milliampères, and finishing with a current of 3 milliampères, so that the total input was from 13 to 15 milliampère-hours, specimens were obtained which gave, on test, deflections of less than 1 cm. A bath of commercial copper sulphate solution, acidulated with sulphuric acid, was used. One interesting phenomenon was observed. When the current was from 2 to 3 milliampères, the deposit was diamagnetic, but by increasing the current up to 6 to 7 milliampères, the deposit was less

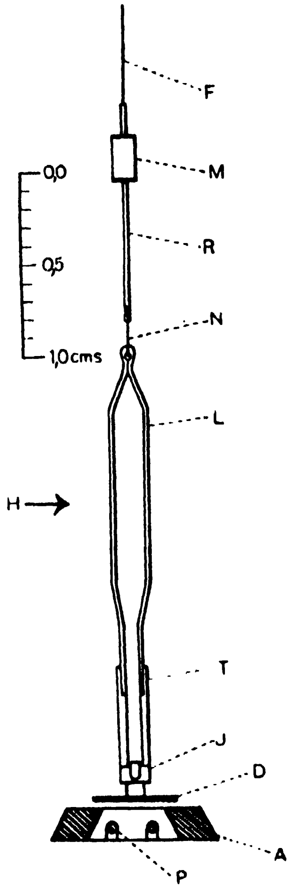
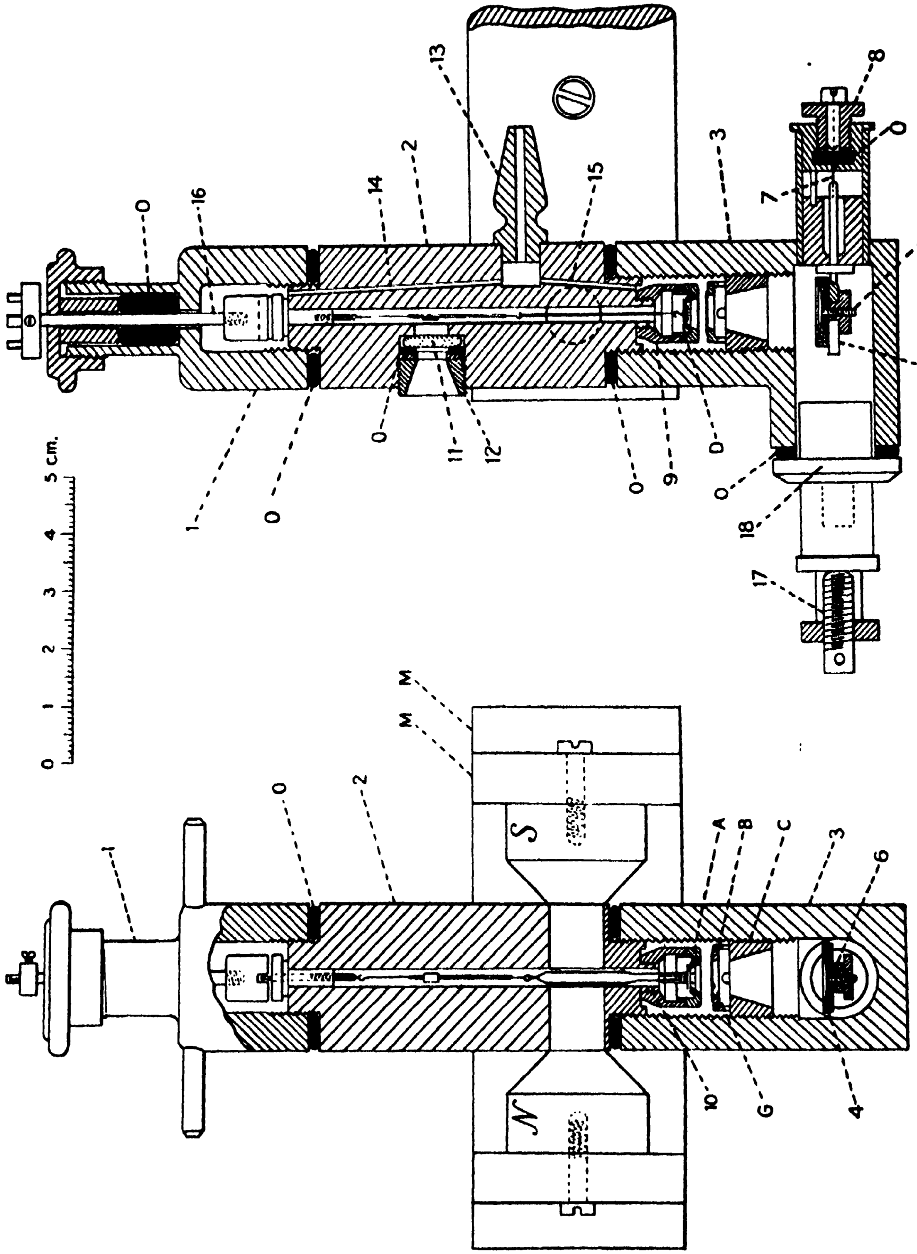


Fig. 1.

---

* Witt, 'Phys. Zeit.,' vol. 21, p. 374 (1920).



diamagnetic, and finally became paramagnetic. This is probably due to impurities in the bath, which are deposited when heavy currents are used.

To the non-magnetic loop was now soldered the thermocouple. The alloys used in its construction were those given by Witt.* The preparation of the strips for the thermocouple was done in the following way: A small portion of the alloy was taken, placed between glass plates, fused and compressed. In such a way sheets of thickness as small as 0·01 mm. could be obtained. In practice, strips of width 0·3 to 0·4 mm. and length 3 to 5 mm. were cut from sheets of thickness 0·03 to 0·05 mm. These were soldered, by means of a soft solder, at one end to the loop at T and at the other end to a small thin silver plate, J. It is necessary to mount the thermocouple to the loop soon after its preparation, because otherwise, if left in the air, after about 3 or 4 hours the loop becomes appreciably paramagnetic, due to the deposit of dust on it. All the work on the loop must be done on a clean glass plate using brass tweezers. The thermocouple cannot be made non-magnetic, and it is essential that it should be placed as far as possible from the magnetic field, as shown in fig. 2.

To the thermocouple, T, is attached, by means of a hook, a silver disc, D, 4·5 mm. diameter, made of a silver sheet 30 mgrm. per cm.^{2} weight, so chosen to absorb the α-particles and to let the β- and γ-radiation through. The disc was attached to the system after the latter was placed in the apparatus. The loop was supported by a silver hook, N, which was sealed into a glass tube, R (diameter 0·2 mm.), to which was attached a mirror, M, dimensions 2 by 1·5 mm., made of a microscope cover-slip. The glass tube was employed by Boys, and serves a double purpose. First, it puts the mirror at a considerable distance from the thermocouple, and thus, owing to the bad conductivity of the glass, shields the couple from any heating effect of the beam of light; and, second, it keeps the mirror well out of the magnetic field.

(ii) *The Source* (fig. 2).—Beneath the plate of the thermo-junction, D, at a distance of 2·5 cm., was placed the source of RaC. This was a brass disc, 5, of diameter 1·2 cm., sometimes plane and sometimes concave, with a radius of curvature 3·5 cm. This was put in a circular receptacle, 4, the edge of which projected slightly above the source, the whole being placed on a fork, 6, which was capable of rotation about a horizontal axis. When the source was rotated 180° from the position shown in fig. 2, no α-particles could strike the plate D beneath the thermo-junction. This method of cutting off the α-rays by rotating through 180° was necessary, because using any other method, for example, a shutter, would disturb the temperature equilibrium, whereas this is not the case on rotation, since the front and back of the source are at the same temperature. A shutter was actually tried, and found to produce an extra deflection of 1 to 2 mm., but rotating the source produced no deflection whatever. The rotation was produced by a device, not shown on the figure, placed outside the apparatus, operated by means of two Bowden cables,* arranged in such a way that the source could be rotated rapidly and without any appreciable vibration of the apparatus.

---

* Witt, *loc. cit.*

Three diaphragms, A, B, C, were placed in the path of the rays. The first, A (see figs. 1 and 2), was made of silver, and was placed at a distance 0·4 mm. from the plate D. The aperture of this diaphragm was 3·5 mm. diameter, whereas that of the plate D was 4·5 mm., so that all the α-particles passing through the diaphragm reached the plate. In this arrangement it is seen that only 1/800 of the α-particles emitted from the source reach the plate. The middle of the diaphragm was crossed by two thin silver wires, P, the purpose of which will be considered later.

The diaphragm, B, was of copper, and was covered by a very thin aluminium foil, G, of stopping-power equal to 1 mm. of air. This foil shielded the thermocouple from the heat radiation from the source, and owing to its small thickness, did not seriously affect the α-particle. The foil G thus separates the lower part of the apparatus from the upper, and so prevents the circulation of any convection currents.

(iii) *The Body of the Apparatus.*—In order to vary the amount of energy lost by the α-particle between the limits required, the pressure of the gas contained within the apparatus had to be varied from 1 cm. of mercury to 3·8 atmospheres. The apparatus has thus to be air-tight, and this was accomplished by means of the rubber rings shown at O.

Since changes in the pressure of the gas are accompanied by changes in temperature, it is necessary to wait some time for the system to reach thermal equilibrium. In order to shorten this time the volume of the gas space must be as small as possible consistent with the fact that the α-particle has to pass through a certain amount of gas, and with the fact that the suspension system occupies a certain volume. As is seen from fig. 2 these ideas were embodied in the construction of the apparatus, the gas space within which was about 5 c.c. The body of the apparatus was made of solid brass so as to have a thermal capacity large compared with that of the enclosed gas.

Another idea embodied in the construction of the apparatus was to shield the thermocouple from the possible variations in temperature occurring within the room. For this purpose we have to avoid any difference in the temperatures of the ends of the thermocouple, JT. *A priori*, it is clear that to keep two points at the same temperatures they must be surrounded by layers of material possessing alternately good and bad thermal conductivity. This was accomplished in the following manner, as shown in fig. 2. The thermo-junction was contained within a cylindrical chamber, 9, made of copper or silver. This chamber is surrounded by air except at the junction, 10, with the body of the apparatus. The second good conductivity layer is the body of the apparatus itself made of brass. The whole of the apparatus including the magnets and the device for rotating the source was contained within a copper box with double walls with water between. Under these conditions the equilibrium of the system was so good that the zero remained constant for periods of several hours.

---

* These cables consist of a spiral steel tube inside of which a steel wire can move, chiefly used on motor cycles.

The body of the apparatus was made of three parts, 1, 2, 3. The lower part, 3, was provided with an opening through which the source was inserted, and which could be rapidly closed by means of the brass stopper, 18, and tightened by means of the single screw, 17. The axle of the source terminated in a pin, 7, passing through a piece of rubber which could be compressed by means of the nut, 8, thus keeping the apparatus tight, and the whole could be rotated from the outside by means of the device already mentioned.

The central portion, 2, was supported by means of the circular pole pieces, N, S, which were screwed to the fixed magnets, M, M. Opposite the mirror of the suspension system an opening was made, which was covered by a plane glass plate, 11, waxed in and kept tight by means of the nut, 12. To the lower central part, 10, different heating chambers, 9, could be attached.

By means of the nozzle, 13, the pressure of the gas within the apparatus could be varied. The gas entering the apparatus had first to pass through the channels, 14 and 15, which were made long and narrow in order that the gas could readily attain the temperature of the apparatus.

The upper portion, 1, consisted of a cap, which was screwed on to the portion, 2, after the suspension system was set up. It contains a device, 16, for adjusting the zero of the suspension system from the outside.

(iv) *The External Apparatus.*—The nozzle, 13, was connected, by means of a rubber tube, in series with a glass reservoir of 300 c.c. capacity, which kept the pressure of the gas within the apparatus fairly constant. The reservoir was wrapped in cotton-wool, this being done because, without it, the temperature variations occurring in the room were transmitted to the reservoir, thus causing changes of pressure of the gas within. This change of pressure is transmitted to the inside of the apparatus which in turn causes changes in temperature therein. By placing one's finger on the reservoir this phenomenon could be easily observed, for a deflection of 1 to 2 cm. was at once produced. On wrapping up the apparatus it was possible completely to get rid of the small movements, of the order of 0·4 to 0·5 mm., of the spot of light on the scale. The apparatus is, in fact, extremely sensitive to changes of pressure, and merely squeezing the rubber tube, which corresponds to a change of pressure of the order of 1/100th of a millimetre, caused a deflection of 40 to 50 cm.

The pressure was observed by means of a simple mercury manometer which was graduated to read directly from + 2·8 to −1 atmospheres. The gas was gradually let out of the apparatus by means of a special tap.

The spot of light was focussed on to a scale, at a distance of 1·3 metres, by means of an eye-piece lens placed close to the opening, 11. The light came from a 100 c.p. pointolite lamp, and in order to absorb the heat in the beam, a glass vessel with parallel walls in which water was circulated, was placed near the lamp. Beyond the glass vessel a short focus lens was placed in order to get a parallel beam of light. This beam now passed through a slit of 1 cm. long and 0·2 mm. wide, a good image of which was observed on the scale. With this arrangement no heating of the apparatus by the light was observed.

Method of Observation.

The source of α-rays was Ra (B+C) obtained by exposing the brass disc, 5, to radium emanation. Its activity, measured by γ-rays, was usually equivalent to from 25 to 45 mgrm. Ra. I am greatly indebted to Mr. G. A. R. Crowe for the preparation of these sources. The source on removal from the emanation was cleaned, measured, and inserted in the apparatus. The apparatus was then adjusted to the proper pressure, or in the case of CO_2 first evacuated to a pressure of 1 mm. the gas then being admitted to the required pressure. After about 10 minutes the apparatus had reached a state of thermal equilibrium and readings were possible. The total time that elapsed between the removal of the source from the emanation and the commencement of observations was from 20 to 25 minutes. In order to save time, a photographic method of observation was arranged, so that it was not necessary to wait until complete thermal equilibrium was attained. The spot of light was projected on a narrow slit behind which a long strip of sensitive paper was moved at a rate of 1 to 1·5 cm. per minute. In attaining thermal equilibrium the spot of light was slowly moving across the slit, and exposing the system to the rays by rotation of the source, photographs as shown in fig. 3 were obtained. From these photographs it is easy to obtain the deflections produced by the α-particles alone. It is only necessary to draw an envelope to the displacement produced by the α-rays and measure the distance from the undisturbed curve.

The photographic method of observation possesses several distinct advantages. It provides an accurate way of recording the time, knowledge of which is necessary in calculating the decay of the source. For this purpose a small electric lamp was placed in front of the slit which, by means of clockwork, gave a flash every minute. These flashes produced the dark lines shown in the photographs.

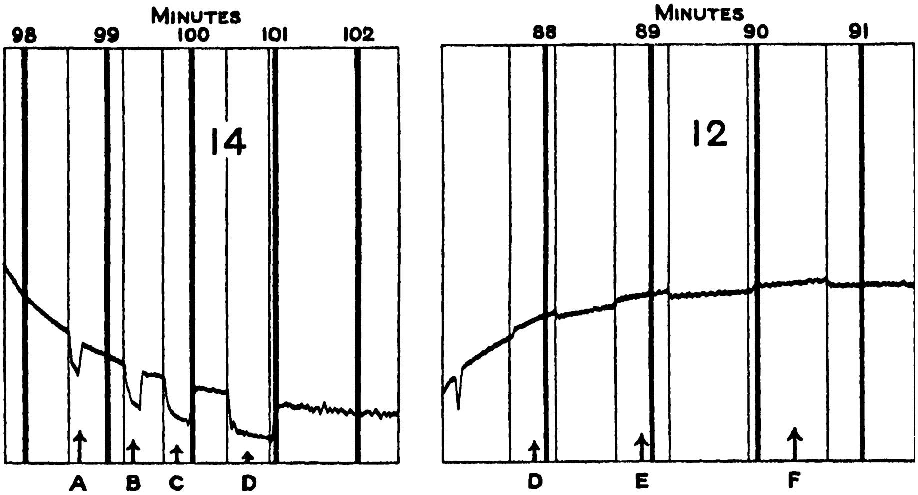


FIG. 3.

The photographic method also permits of the mean value of the observations being observed in spite of the fluctuations produced by vibration of the apparatus.

After obtaining three or four readings, the pressure was reduced to a new value and the procedure repeated. After changing the pressure, the time to reach approximate equilibrium and take the next observation was from 4 to 6 minutes. Hence if the total time taken from the removal of the source from the emanation to the last observation is 100 to 110 minutes, it is possible to obtain from fifteen to nineteen points on the curve, for each of which some two to four observations could be made.

In all the experiments the pressure was reduced to obtain successive readings. The lowest pressure was not less than 1 to 2 cm., for with lower pressures the charge acquired by the suspension system cannot be readily lost by ionisation in the surrounding space, and the electrostatic couple thus produced causes inaccuracies in the readings. The gases used were air and carbon dioxide taken from a cylinder. The accuracy with which the photographs obtained could be measured up, using a magnifying glass, was 1/10 mm. The observed deflections were then corrected for the decay of the source. The distance between the plate D and the source was measured by means of a cathetometer.

Knowing the temperature and the pressure within the apparatus, the amount of gas traversed by the α-particle before it reaches the plate D can be readily calculated. These distances were then expressed in terms of pressure 760 mm. and temperature 15° C., and were plotted as abscissæ against the observed deflections as ordinates. The relations obtained are shown in Diagrams I, II, III. If the deflection is proportional to the kinetic energy

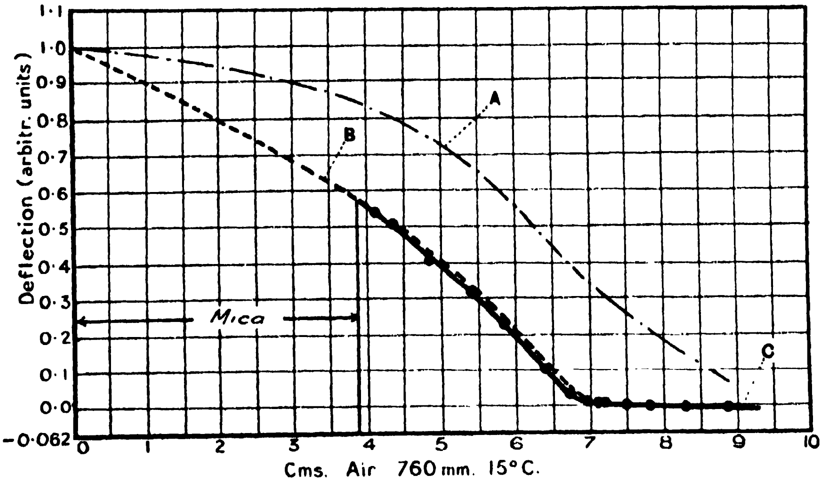


DIAGRAM I.

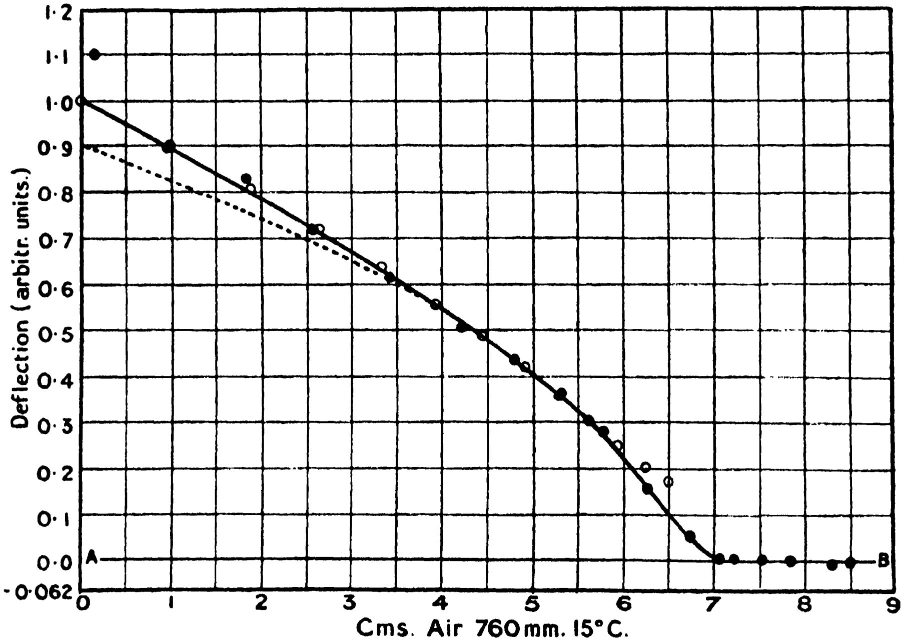


DIAGRAM II.—Energy distribution curve in air. ○ ... Marsden and Taylor ; • ... Kapitza ; – – – – – ... Henderson.

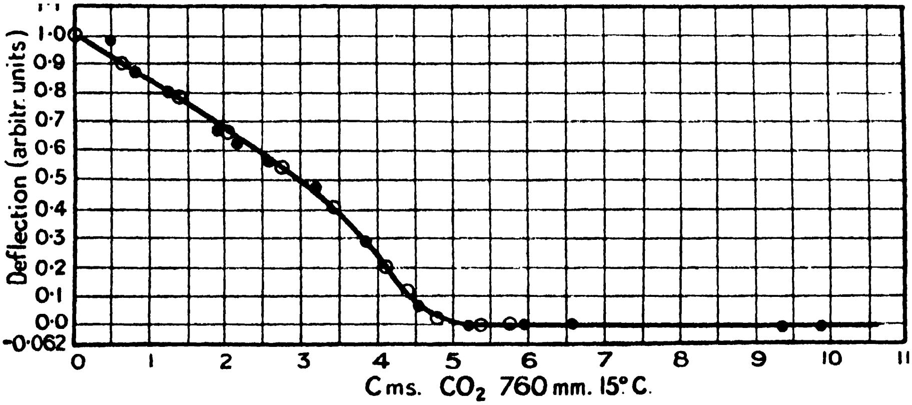


DIAGRAM III.—Energy distribution curves in CO_2 and air. ● ... CO_2 ; ○ ... Air.

of the particle, then the curves will represent the distribution of energy of the α-particle at different parts of its range.

Stray Effects.

Before attempting an interpretation of the curves obtained it is necessary to examine the stray effects produced in such a sensitive apparatus. We will first consider any occasional effects and finally the systematic stray effects.

We have considered the occasional stray effects produced by variations in temperature in the room, and the heating effect of the beam of light; but, due to the precautions previously mentioned, these effects caused no appreciable error, and never interfered with the observations. Some inconvenience was caused by variation of the pressure within the apparatus, due to small leaks of the order of 1 mm. in 3 or 4 hours. According as the pressure within the apparatus was greater or less than atmospheric the zero of the suspension system had a new value, but remained quite steady. It was feared that this variation in the zero might produce a variation in the sensitiveness of the apparatus. On the other hand, because the suspension system was not entirely non-magnetic, it was feared that the deflections produced might not be proportional to the heating produced in the plate D by the absorption of the α-particles. To examine these points the following experiment was performed. Beneath the thermocouple a very thin platinum wire was placed through which different currents were sent. The results obtained showed that the deflection was quite proportional to the square of the current passing, and that the sensitiveness of the apparatus was quite independent of the pressure of the gas. It is thus seen that these effects produce no error in the readings obtained.

On account of the small size of the suspension system, and the consequent unsteadiness produced in it by vibrations of the building, it was always necessary, in order to obtain good results, to work during the evening.

The first systematic stray effect is due to the fact that only a small fraction (about 1/800) of the total number of α-rays emitted by the source reached the plate D. The rest of the rays lose their energy in heating the body of the apparatus. Since this amount of energy is considerable, if no special precautions were taken the stray effect would be large. The diaphragms, B, C, were entirely for the purpose of protecting the heating chamber, 9, from this effect. To obtain some idea of the order of magnitude of the effect, the following experiment was made. The diaphragms B and C were removed and a loop of thin wire placed some distance below the heating chamber, through which a current was passed of such a magnitude as to produce approximately the same heating as that produced by the α-rays. It appears that the effect does exist, but it does not, on account of the large thermal capacity of the apparatus, reach its maximum value until after 5 minutes. The time taken for the suspension system to reach its maximum deflection, when the α-rays are absorbed in the plate D, is, however, from 5 to 10 seconds. On account of this difference in time we can appreciate the magnitude of the stray effect. In fig. 3, photograph 14, four observations are noted, A, B, C, D, which correspond to exposures of 5, 10, 20, 30 seconds. No appreciable differences in the deflections are observed. The final small kicks in the deflections are due to the fact that in the act of rotation one edge of the source comes closer to the plate.

A further systematic stray effect is due to the fact that the α-particles, before striking the plate D, heat the surrounding air. This effect may change entirely the nature of the phenomenon, unless proper precautions are taken. The broken curve, A, Diagram I, represents the results obtained in a previous experiment without the use of diaphragms. This curve shows that the deflections towards the end of the range of the α-particles are greater than one would expect, and this is due to the fact that the α-particles heat the air in the neighbourhood of the edges of the plate. In addition, the curve extends beyond the end of the range of the α-particles, and this is due to the heating of the air immediately beneath the plate. The diaphragm A was inserted in order to avoid these two stray effects. As already noted, the diameter of the aperture in A was less than that of the plate D, and the relative positions of A and D prevented any α-particles from getting near the edges of the plate. To avoid the heating of the air beneath the plate, two fine silver wires, P, crossed the diaphragm at a distance 0·7 mm. from the plate. On account of the high thermal conductivity of these wires, the heat produced in the air was prevented from reaching the plate, and, instead, was conducted to the body of the heating chamber. With these precautions the magnitude of these stray effects was reduced below the experimental error, as can be seen from the results of the following experiment:—

The source was covered by a mica sheet of stopping-power equal to 3·75 cm. of air, and the curve C obtained. The broken curve B was obtained without the use of the mica sheet. In the former case the whole range of the α-particle was obtained with a pressure just one-half that in the second case. Consequently, the heat produced in layers of air of equal thickness below the plate should be twice as great in the second case as in the first, and so the magnitude of these stray effects ought to have the same variation. The good agreement of the curves B and C shows that the effects could not be detected.

The final systematic stray effect is due to the absorption of the β-rays, and will be considered later in a special section.

Sensitiveness and Accuracy of the Method.

The sensitiveness of the apparatus was considerably reduced by the use of the diaphragm, A, and the silver wires, P. This is easy to understand. From equation (i) it is seen that the deflection, θ, is proportional to the increase in temperature, ΔT, of the thermo-junction. This increase in temperature is given by

$$\Delta T = \frac{Q}{K_a + K_j},$$

where Q is the amount of heat given to the plate per second, K_a, K_j are respectively the whole thermal conductivities of the air between the plate, D, and the heating chamber, 9, and the strips forming the thermo-junction. In practice K_j is small compared with K_a. Putting the diaphragm, A, nearer the plate, K_a is considerably increased, causing a consequent reduction in the sensitiveness of the apparatus. The sensitiveness of the apparatus was found to be about four times greater when the diaphragm, A, was omitted. It was possible to overcome this loss of sensitiveness in the following manner:— From equation (i) it is seen that the deflection produced is almost inversely proportional to the resistance of the thermocouple (the resistance of the loop is small), and making the strips of the thermocouple short and thick, r, is diminished. However, this increases K_j, since $K_j \propto 1/r$. Diminishing r increases the sensitiveness of the apparatus, but increasing K_j does not appreciably affect ΔT, since K_j is always small compared with K_a. We thus see that the thermo-junction should not be made of thin strips. However, this method of overcoming the difficulty has two disadvantages which, in practice, limit the sensitiveness of the apparatus.

First, we see from equation (i) that the deflection, produced by the stray effects within the apparatus, is proportional to the term .

$$B\frac{SH}{r\mu}$$

and is independent of K_j and K_a. Decreasing r simply increases the above term, and we obtain finally a system which is very sensitive to stray effects, and which takes a long time to attain thermal equilibrium.

Second, making the thermo-junction of thick strips, it was difficult to produce a non-magnetic suspension system.

The sensitiveness of the apparatus in the experiments was as follows:— The maximum deflection produced per milligram of RaC on the source was 1·4 mm. in CO_2 and 0·95 mm. in air. The deflection in air is less on account of the higher thermal conductivity. Using 40 mgrm. of RaC as the source, the maximum deflections produced were 5·6 cm. in CO_2 and 4 cm. in air. On the photographs a deflection of 1/10th of a millimetre could be measured. Taking into account the decay of the source, after removal from the emanation, this shows that it is possible to measure about 1/300th of the initial energy of the α-particle. As the readings were obtained working back from the end of the range the deflections observed were generally between 5 and 10 mm. Thus the accuracy of measurement of a single observation was about 2 per cent. As each point was the mean of several observations it is probable that the actual accuracy is better than this. The error due to the fact that some of the α-particles had to travel longer distances to the plate, was reduced to about 1 per cent. by using a concave source.

Discussion of Results.

The full curve, Diagram II, represents one of the characteristic curves obtained for air. The ordinates of the curve decrease continuously until a range of about 6·95 cm. is reached, afterwards remaining constant. This final constant deflection, which is about 5·8 per cent. of the total initial deflection, is due to the β-rays as discussed in the next section. To separate the effects of the β-rays from those of the α-rays, it is only necessary to draw the horizontal line, AB, which is the prolongation of the observed β-ray effect. This is probably quite legitimate, since the β-rays are only slightly absorbed in the air-space traversed by the α-rays. Thus the curve referred to AB as axis will represent the true energy distribution of the α-particles. From these considerations it is seen that the α-particle has no energy beyond the end of its range as determined by ionisation methods. A similar curve was obtained for CO_2, as shown in Diagram III. Allowing for the β-ray effect the range for the α-particles is found to be 4·75 cm., which is in good agreement with previous determinations. Changing the scale of the abscissæ for the curve obtained with air in the ratio 4·75 to 6·95, the curves for air and CO_2 (see Diagram III) were found to coincide very nearly. This result shows that the loss of energy of the α-ray in CO_2 is governed by the same laws as in air.

An attempt was now made to compare the curves obtained with those obtained by the ionisation method. If the α-particle produces N-ions per centimetre path, then in a distance dx it will produce N dx-ions, and if the energy to create a single ion is E, the energy required in this production will be

$$\mathrm{EN}\,dx.$$

If W is the total energy of the α-particle, W_x its energy after travelling a distance x, then if all the energy lost is due to ionisation,

$$W_x = W - \int_0^x \mathrm{EN}\,dx.$$

If we assume that the same amount of energy is required to produce an ion at all points of the range, then we obtain

$$W_x = W - E\int_0^x N\,dx.$$

The dependence of N on x is given by the ordinary ionisation curves, the best being those obtained by Henderson* which were used in the evaluation of the integral. The curve calculated in this way is shown by the broken line on Diagram II. It is seen that after 4 cm. of the range, the curves are in very good agreement. This shows that towards the end of the range the average energy required for the production of each ion is constant, and we can assume that the loss of energy is proportional to the ionisation produced. Near the beginning of the range, the Henderson curve shows a marked departure from the observed curve, indicating that the energy required to produce an ion is greater for the swifter rays. We can thus say that the energy required for the production of an ion by an α-ray of range 7 cm. is, on the average, about 10 per cent. greater than that required by an α-ray of range 3·5 cm. This leads to the conclusion that the α-rays with high velocities are better able to eject electrons from inner orbits of the atom, which process would require more energy than the ejection of electrons from the outer orbits. This confirms the theory advanced by Bohr† of the passage of α-rays through matter.

---

* G. H. Henderson, 'Phil. Mag.,' vol. 42, p. 538 (1921).

† N. Bohr, *loc. cit.*

A further comparison can be drawn between the present results and those obtained by Marsden and Taylor.* These latter are shown by open circles on Diagram II. These points are obtained from Table III, column 5, p. 448, of Marsden and Taylor's paper, which they state are probably the most exact values, and it is seen that, up to a range of 6 cm., the agreement between their curve and mine in Diagram II is very good. Beyond 6 cm. they no longer agree, the present experiments giving lower values for the energy. In the following Table are compared the numerical values obtained by Marsden and Taylor and in the present work.

	Range in air......	5·78 cm.	6·26 cm.	6·50 cm.	6·72 cm.	7·07 cm.
Kapitza	V	0·522	0·400	0·316	0·244	Less than 0·084
	V^2	0·273	0·160	0·100	0·059	Less than 0·007
Marsden and Taylor	V	0·527	0·450	0·415		
	V^2	0·278	0·202	0·172		

where V, V^2 are expressed in terms of initial values. The disagreement between the values of V^2 when the range is 6·50 cm. is 72 per cent., and this cannot be explained by experimental errors. A complete explanation of this disagreement is not easy, but the following considerations may help to throw further light on the problem.

First, in the present experiments, the average energy of the α-rays in a beam was measured, and towards the end of the range it is possible that the increased scattering may affect the result. On account of this scattering some α-particles may be deflected away from the plate, causing a diminution in the total number which should fall on the plate. This diminution should depend on the distance from the plate at which scattering becomes appreciable, and should thus be quite different in the case when the rays pass direct through air to the case already mentioned, when they first pass through a mica sheet of stopping-power equal to 3·75 cm. of air. The good agreement between the curves B, C, Diagram I, shows, however, that any difference produced is too small to measure, so that scattering considerations will not explain the discrepancy.

Second, we have to consider the lack of homogeneity in the beam. On theoretical grounds, the loss of energy of an α-particle is due to interaction with a very large number of atoms, and any deviation from the statistical mean should thus be very small. We are thus led to anticipate that the beam should be fairly homogeneous. Practically, however, it is a difficult matter to obtain a homogeneous beam of α-rays, owing to the lack of uniformity in the thickness of the material used to absorb them. From the energy curve, shown in Diagram II, it is seen that, towards the end of the range, from 6 to 6·5 cm., an error of 4 per cent. in the range causes an error of 100 per cent. in the energy. In the experiments of Marsden and Taylor, as pointed out by the authors, the scintillations towards the end of the range became fainter and fainter, so that it was impossible to detect α-particles with energies less than 0·172 of the initial energy. Thus it seems probable that, in the α-ray beam used by Marsden and Taylor, there were always some rays with velocities greater than the mean, due to the lack of uniformity in the thickness of the absorbing material, and on account of the loss of sensitiveness of the apparatus, only these swift particles, which constitute a small fraction of the total number, could be observed. The large variations in the measurements of the speeds of the α-rays towards the end of the range, using as stopping materials mica, aluminium and gold, support this view.

---

* E. Marsden and T. S. Taylor, *loc. cit.*

Further, Marsden and Taylor have pointed out that the inaccuracies in their work may be due to the fact that some of the α-particles are singly charged owing to recombination. There is some reason to suppose that this does happen, and perhaps the α-particle gains and loses electrons several times during its passage through matter, and towards the end of the range the number of recombined particles is increased. If we suppose that the majority of α-particles, whose energy is less than 0·172 of the initial energy, are singly charged, this will account for the abrupt limit of the observations of Marsden and Taylor.

In the present experiment the α-ray beam should be fairly homogeneous, on account of the fact that the stopping material was a gas. Any lack of homogeneity in the α-ray beam arising from variations in velocity or from differences in the charge carried will, however, produce no effect, since a mean value was measured, and so the velocities observed should agree most closely with those of an "ideal" α-ray beam.

Finally, it is interesting to compare Marsden and Taylor's curve with that obtained from the ionisation curve in the manner already described. We find that the curves fail to agree beyond a range of 6 cm. in such a way that, at 6·5 cm., the energy required for the production of an ion is twice as great as that required at 6 cm. It is difficult to account for this variation on theoretical grounds.

The Heating Effect of the β-Rays.

In the previous section, we have assumed that the deflection obtained after the end of the range of the α-rays was due to the heating effect of the β-rays, and the line AB was drawn in order to separate the two effects. These considerations are so important to the correct interpretations of the results, that special experiments were carried out to justify the assumption. It is certain that this deflection after the end of the range cannot be due to the α-rays. On Diagram III, which represents the curve for CO_2, two points are obtained at distances 9·6 and 10 cm., to produce which the α particle would, after the end of its range, have to pass through a layer of CO_2 of stopping power greater than that of the plate D of the thermo-junction. But the deflections corresponding to these two points are only about 20 per cent. less than those immediately after the end of the range. This shows that the effect produced after the end of the range is due to a radiation of greater penetrating power than the α-radiation. For a more detailed examination of the effect the following arrangement was set up.

A small thin-walled glass bulb, diameter 3 mm., shown at 2 in fig. 4, was filled with 46 mgrm. of emanation, the thickness of the walls being just sufficient to absorb the α-rays. The bulb was fixed by means of a little wax, 3, into a brass receptacle, 1. The opening at the top of the receptacle could be covered by successive layers of aluminium foil, which could be fastened down by means of the cover, 5. This arrangement now replaced the source on the fork, 6, fig. 2. Rotating this arrangement, the β-rays could be allowed to fall on the plate at will. In this way deflections were obtained of the same order of magnitude as those found in the previous experiments after the end of the range of the α-particles. The apparatus was now divided into two parts by means of a thin aluminium sheet of thickness 0·045 mm., placed beneath the diaphragm, C, fig. 2. The deflection was reduced but did not disappear, which shows that the effect could not be due to any heat radiated from the source, or to any convection currents in the gas.

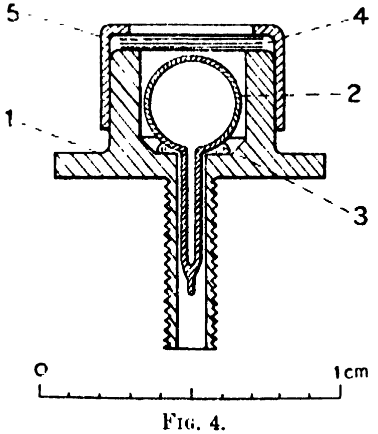


FIG. 4.

Afterwards a disc of lead, thickness 0·11 mm., was attached to the plate, D. The deflection now produced was about 1·35 times greater, apparently due to the greater absorption of the β-rays. Removing the diaphragm, A, the deflection was considerably increased and it was possible to obtain a curve for the absorption of β-rays in aluminium. By altering the number of foils, 4 (fig. 4), each of thickness 0·045 mm., Diagram IV was obtained. From this curve it is possible to calculate the coefficient of absorption, μ, for β-rays in aluminium. We find, at the beginning of the curve, $\mu = 56$, and at the end $\mu = 26$, which is in good agreement with the limits, $\mu = 65$ to $\mu = 13$, determined by ionisation methods.

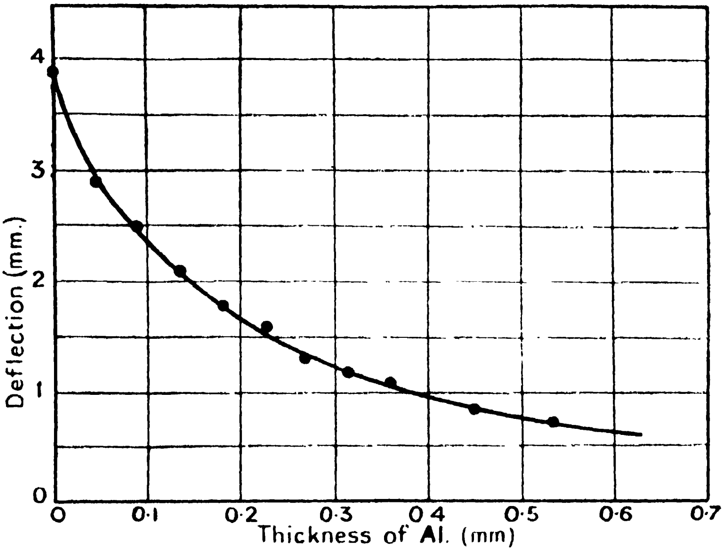


DIAGRAM IV.—Absorption of β-Rays in Aluminium.

All these experiments confirm the original assumption. From Diagram I and the previous experiment a very rough estimate can be formed of the energy carried by the β-rays compared with that carried by the α-rays. Curve C, Diagram I, shows that the β-ray effect is about 6·2 per cent. of the α-ray effect. Taking into account the increased absorption in the lead plate this value becomes 8·4 per cent. The value obtained by Rutherford and Robinson* is 10·9 per cent. The difference of 25 per cent. between these values cannot be regarded as established, because the present method is open to two objections. First, it is quite probable that the lead plate does not absorb all the β-rays, but this is not sufficient to account for such a large difference. Second, on account of the scattering of the β-rays by the plate D a portion only of the energy in the β-ray beam is converted into heat. On the other hand, β-rays scattered from the walls of the apparatus and falling on the plate will compensate for this to some extent. The complicated shape of the interior of the apparatus renders it difficult to make an exact estimate of the magnitude of these scattering effects. Further experiments, using an apparatus of more suitable shape, are needed to settle the point.

---

* E. Rutherford and H. Robinson, *loc. cit.*

Analogy to the Ionisation Curves.

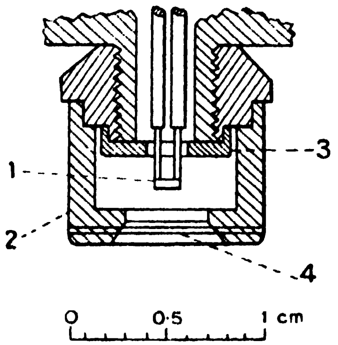


Fig. 5.

The magnitude of the stray effect due to the heating of the air in the neighbourhood of the thermocouple led to the idea of measuring the heating effect at different parts of the range, thus providing an analogy with the ordinary ionisation curve. For this purpose the heating chamber 9, fig. 2, was replaced by a special type shown in fig. 5. The actual chamber, 2, was only 3 mm. deep, containing an opening 4·5 mm. diameter crossed by two silver wires, 4. The thermocouple, 1, projected into the chamber through an opening, 2 mm. diameter, in the cover, 3. In this way it was possible to isolate a small quantity of gas and to measure the heating therein produced. Altering the pressure in the apparatus changes both the mass of gas contained in the heating chamber and the amount of gas through which the α-particle has first to pass. Working in the same way as before, curve 1 shown in Diagram V was obtained. To obtain from this the analogy to the ionisation curve it is necessary to apply some corrections. First, some α-particles struck the thermocouple directly, thus producing extra heating. It is possible to separate this stray effect from the real effect in the following manner. The point A on the curve corresponds to the full exhaustion of the apparatus and the whole deflection is due to the α-rays which strike the thermocouple. If we now draw from A the energy curve, 3, similar to that previously obtained, the difference between the ordinates will represent the energy due to the heating of the air in the heating chamber. The correctness of this interpretation was tested experimentally in the following way:—

A small silver plate, of dimensions 2 mm. by 1 mm., was supported on the silver wires, 4, thus partly shielding the thermocouple from the direct α-ray beam. In this way the curve 4 was obtained, and it is seen that the initial deflection is considerably reduced. These results cannot be used in further calculation, due to the fact that the sensitiveness of the apparatus was reduced by one-half, because the plate stopped some α-particles which would have heated the gas surrounding the thermocouple. Correcting curve 1 by means of curve 3, we can thence obtain the deflection per unit mass of heated gas by dividing the ordinate by the density. In this way the continuous curve in Diagram VI was obtained.

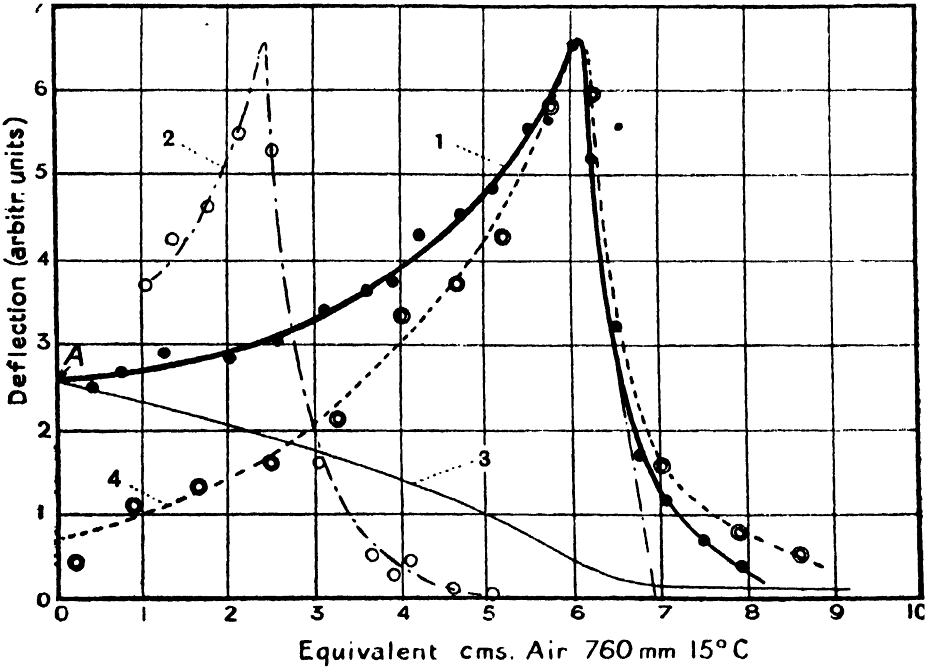


Diagram V.

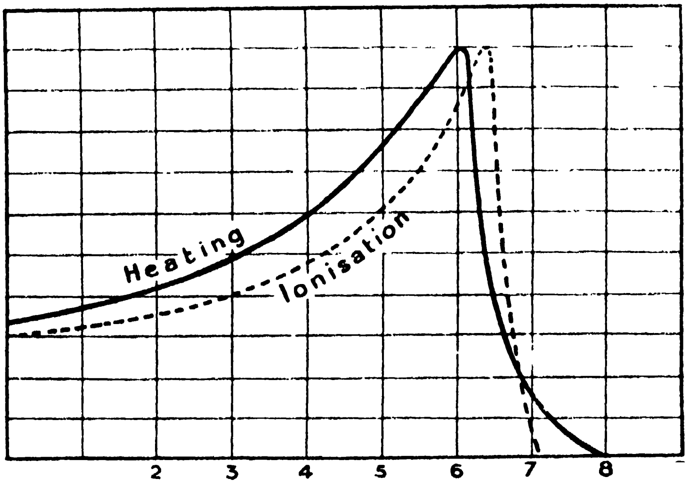


Diagram VI.—Analogy to ionisation curve. ——— ... heating curve
- - - - ... ionisation curve.

Before this curve can be compared with the ionisation curve, it is necessary to establish the end of the range. This was done in the following way: A mica sheet of stopping-power equal to 3·75 cm. of air was placed over the source and the curve 2 obtained. The slope of this curve beyond the maximum is very similar to that of curve 1. The distance between these two portions of the curves is equivalent to 3·75 cm. stopping-power of air, and thus we can deduce the scale of abscissæ, fixing the range at 6·95 cm. In this way, taking into account the pressure, it is found that the end of the range of the α-particles is at a distance 0·8 mm. below the thermo-junction. This result is to be expected, because of the high thermal conductivity of the thermocouple and the opening, 4, in the chamber. Having fixed the end of the range, we can now compare the curve with that obtained by Henderson.* This is represented by the broken curve in Diagram VI.

The curves appear to possess the same general characteristics but do not fully agree. This is due to the fact that, in the present experiments, observations were not made on a single point, but a mean taken over a definite region, *a*, about 0·12 of the total distance between the source and the thermocouple. When an allowance is made for this distribution, the shapes of the calculated and observed curves are very similar.

Further speculation in regard to the curve obtained is of little purpose owing to the inaccuracies in the experiment. These inaccuracies are due (i) to the fact that the chamber measured the heating over a considerable portion of the range; (ii) to the fact that the deflections obtained, using CO_2, were always small, of the order 4 to 5 mm., giving an accuracy of 2 to 3 per cent.; (iii) to the fact that, in order to obtain the curve, several operations have to be performed which probably considerably increase the error.

It is interesting to note the sensitiveness of the apparatus as used in obtaining the above curve. Fig. 3, curve 12, represents one of the initial points obtained on curve 1, Diagram V. The deflections are due mainly to the α-particles which strike the thermocouple directly. The number of α-particles producing this deflection of 1 mm. was approximately calculated to be 2000 to 3000 per second. This corresponds to a rate of heating of the order of 10^{-9} calories per second.

Summary.

(1) The construction of a sensitive radiomicrometer, and its application to the measurement of the distribution of energy in an α-ray beam is described.

(2) With the apparatus, the energy distribution in an α-ray beam in its

---

* G. H. Henderson, *loc. cit.*

passage through air and CO_2 has been determined, and it has been shown that, within the limits of the experimental error (1 to 2 per cent.), the same law governs the distribution in the two gases.

(3) Immediately beyond the end of the range, the α-particles have an energy not greater than 0·7 per cent. of the initial energy.

(4) A comparison of the energy distribution curve with that obtained from the ordinary ionisation curve, shows that α-particles with high velocities on an average expend more energy in the production of an ion than those moving with low velocities.

(5) A comparison of the energy distribution curve with that obtained by the measurement of the velocities of the α-particles by deflection in a magnetic field, shows that, up to a range of 6 cm., the curves are in good agreement, but that beyond this point the energy curve gives distinctly smaller values. An attempt to explain this discrepancy has been made.

(6) The heating effect of the β-rays was detected, and by means of a modified apparatus the energy absorption curve in aluminium was determined.

(7) The heating produced in CO_2 over different portions of the range was measured, and in this way curves very similar to the ordinary ionisation curves were obtained.

The passage of α-particles through solid bodies will be described in another paper.

It is a great pleasure to thank Sir Ernest Rutherford for permitting the work to be carried out in the Cavendish Laboratory, and for his active interest and kind advice during its progress. I am also deeply indebted to my teacher, Prof. A. F. Ioffé, for providing the opportunity to work in Cambridge. My best thanks are due to Mr. M. H. Belz for help in the composition of the paper.

The Relation between the Evolution of Heat and the Supply of Energy during the Passage of an Electric Discharge through Hydrogen.

By J. Keith Roberts, M.Sc., Trinity College, Cambridge; Exhibition of 1851 Scholar of the University of Melbourne.

(Communicated by Prof. Sir E. Rutherford, F.R.S. Received June 21, 1922.)

I. INTRODUCTION.

In this paper an account is given of measurements of the heat evolved by the passage of a known current under a known potential through a hydrogen discharge tube. The experiment was suggested by Sir Ernest Rutherford as a possible test of the hypothesis in which he supposed that a much closer combination between an electron and a hydrogen nucleus than that existing in the hydrogen atom might occur under suitable conditions.* The experiments of Aston, which show that the masses of all atoms except hydrogen are whole numbers, indicate that the close binding of electrons and positive nuclei causes a decrease of mass, owing to the interaction of the fields. Such a decrease of mass might be expected to occur if the supposed close combination of a single hydrogen nucleus and an electron took place. We should expect this decrease of mass to be accompanied by an evolution of energy, since a mass m has energy $\frac{1}{2}mc^2$ associated with it, c being the velocity of light.

In order to account for the fact that the energy radiated by giant stars is too great to be supplied by the loss of gravitational potential energy by contraction, it has been suggested† that atomic changes of this nature occur in these stars. If 1 per cent. of the mass of a giant star were converted from hydrogen to helium the energy evolved would be sufficient to supply the star with energy for $1 \cdot 5 \times 10^7$ years.

In such a star violent collisions will occur between electrons and hydrogen nuclei owing to the high temperature. Such collisions will also occur in a discharge tube containing hydrogen. If in this case these collisions give rise to the changes mentioned, it might be expected that more heat would be evolved from the discharge tube than could be accounted for by the supply of electrical energy as measured by the product of the current and the potential difference.

---

* 'Roy. Soc. Proc.,' A, vol. 97, p. 396 (1920).

† Perrin, 'Annal. de Phys.,' vol. 11, p. 90 (1919); Sommerfeld, 'Atombau und Spektral-Linien,' p. 538 (1919); Eddington, 'Nature,' Sept. 2, p. 19 (1920).

To test whether this was the case, a discharge tube was placed in a calorimeter and a comparison made between the heat produced and the energy supplied. It was found that these two quantities agreed to less than $\frac{1}{2}$ per cent., which was within the error of the experiment.

To appreciate the meaning of this negative result, it is necessary to consider in such detail as is possible the conditions prevailing in the discharge tube. The energy of translation of an electron after falling through one volt is $1{\cdot}59 \times 10^{-12}$ ergs, since its charge is $4{\cdot}77 \times 10^{-10}$ e.s.u., and one volt is equal to 1/300 e.s.u. The mean energy of translation of a molecule of a gas at absolute temperature T is equal to $3/2$ RT, where $R = 1{\cdot}37 \times 10^{-16}$. The electron after falling through one volt therefore has translational energy equal to the mean translational energy of a gas molecule at temperature T given by

$$T = \frac{2 \times 1{\cdot}59 \times 10^{-12}}{3 \times 1{\cdot}37 \times 10^{-16}} = 7{,}730^\circ.$$

We must now consider the mean free path of the electron. To do this an estimate must be made of the radius of the sphere of action of the hydrogen atom. This will be of the order 10^{-8} cm., which is sufficiently accurate for the present purpose. Taking the number of atoms of hydrogen in a cubic centimetre of gas at atmospheric pressure as $2 \times 2{\cdot}7 \times 10^{19}$, the mean free path, λ, of an electron which is both so small that it can be regarded as a point, and is moving so fast that the hydrogen molecules may be assumed at rest, is given by

$$\lambda = \frac{760}{10^{-16} \times 2 \times 2{\cdot}7 \times 10^{19} \times p} = \frac{0{\cdot}04}{p},$$

where p is the pressure in millimetres of mercury.

The conditions under which the discharge tube was run were such that there was no positive column. In this case nearly the whole drop of potential occurs in the cathode dark space, and the potential gradient is uniform within that region.* If we assume that at each collision an electron loses the whole of its energy, the minimum value possible for the average energy with which an electron leaves the dark space can be obtained by finding the potential drop in a free path and calculating the energy gained by an electron in falling through this potential.

We cannot say what happens to the electrons after they leave the dark space. The only information that this calculation gives us is that electrons are projected from the dark space into the gas in the remainder of the tube, with energies of this order and greater energies. In the experiments carried out the conditions lay between the limits indicated in the following table:—

---

* Aston, 'Roy. Soc. Proc.,' A, vol. 84, p. 530 (1911).

Pressure (mm. of mercury).	Length of dark space (cm. L.).	Mean free path of electron (calculated) (cm. λ).	Number of free paths in dark space ($L/\lambda = n$).	Total voltage (V).	Voltage per mean free path (V/n).	Temperature corresponding to energy of electrons after falling through (V/n).
						°
0·46	0·9	0·09	10	870	87	670,000
0·13	2·5	0·30	8·3	2720	330	2,600,000

The average energy with which the electrons are projected through the gas probably corresponds to considerably higher temperatures than these. The upper limit to this energy is fixed by the total voltage drop. This would correspond to temperatures of 6,700,000° and 21,000,000° respectively.

The changes which give rise to the large evolution of energy in giant stars certainly take place in those of type M. For these, according to Eddington's theory of the internal constitution of stars,* the maximum temperature at the centre is 1,400,000°. For giant stars of type F, in which the changes still occur, the maximum temperature is 6,590,000°. These figures show that the energies of the electrons in the discharge tube were such that the effect might have been expected to take place.

The order of effect to be expected cannot be estimated. The energy radiated by a giant star amounts to roughly 200 ergs per second per gramme mass of star. We do not know how much of this mass is hydrogen. The mass of hydrogen present in the discharge tube of volume 70 c.c. would be 0·0000042 grm. at a pressure of 0·5 mm. of mercury. Assuming that the star consists wholly of hydrogen, and that the same figure, namely, 200 ergs per second per gramme, which applies to all giant stars, also applies to the discharge tube, the energy evolved by these changes would be 8×10^{-4} ergs per second. This would be about 10^{-11} of the heat evolved by the passage of the discharge. We see, therefore, that the negative result of this experiment does not in any way affect the hypothesis put forward by Eddington as to the origin of stellar energy.

Another point that should be mentioned is that, if the close combination of an electron and a hydrogen nucleus did occur, the energy evolved might not appear directly as heat, but might come out in the form of a very penetrating γ-ray, which would escape detection in ordinary experiments.

---

* 'Zeits. für Physik,' vol. 7, p. 351 (1921).

II. EXPERIMENTAL DETAILS.

Discharge Tube and Gas Supply.

The apparatus consisted of a cylindrical discharge tube 19 cm. long and 2·5 cm. in diameter, immersed in paraffin oil contained in a large Dewar flask 8 cm. in diameter and about 30 cm. long. The electrodes were aluminium discs 2 cm. in diameter. The tube was exhausted by a Gaede rotating mercury pump. In some of the experiments the gas was absorbed, and in order to keep the conditions in the tube constant, hydrogen prepared by electrolysis of dilute sulphuric acid was allowed to leak into it through a capillary from a reservoir, the pressure in which was adjusted to a suitable value. The discharge tube itself could not be seen, and, in order to be able to observe under what conditions the discharge was running, a similar tube was connected to the tube leading to it. Arrangements were made so that this outside discharge tube could be substituted for the one inside the flask in the electrical circuit. This tube was used to measure the length of the dark space. For the same pressure, the two tubes were practically identical as regards the relation between current and voltage.

The flask was closed by a cylindrical wooden block, fitting over its mouth, which carried the discharge tube, stirrer, etc., which were inside it.

Thermostat.

Preliminary experiments showed that one of the serious difficulties of the experiment was the relatively large value of the heat losses. This was due to the fact that the power which can be put through a discharge tube is so small in proportion to the volume it occupies. If one tries to increase the power beyond certain limits, the discharge becomes very unsteady. This fact made it necessary that the heat losses should be definite and determinable. In order that this should be the case, the flask was immersed in an electrically heated thermostat with a toluene mercury control. The temperature of this thermostat was constant to less than 1/100° during an experiment. This temperature, which determined the most important heat loss, was read on a Beckmann thermometer.

The air temperature, which determines the loss through the wooden top of the flask, was read on an ordinary mercury in glass thermometer.

Measurement of Rise of Temperature.

The rise of temperature of the oil was measured by five naked copper-constantan thermo-junctions in series. The cold junctions inside a glass tube were immersed in a mixture of ice and distilled water contained in a

Dewar flask. Each lead to the junctions, both those in the oil and the cold junctions passed through a separate glass tube to ensure that no insulation leaks should occur.

The oil was stirred by a metal stirrer carried on a glass rod. The stirrer just cleared the walls of the flask, and was proved to be efficient by connecting two opposed thermo-junctions, one at the top and the other at the bottom of the flask, straight through the galvanometer. A revolution counter was attached to the stirrer.

Arrangement of Circuits.

The electrical arrangements are shown in fig. 1. The discharge tube D and a megohm, M, were connected in parallel. The milliammeters A_1 and A_2 were used to determine the currents in these two parts of the circuit. The current through the megohm determines the voltage drop over the tube. The resistances R_1 and R_2 were used to calibrate A_1 and A_2.

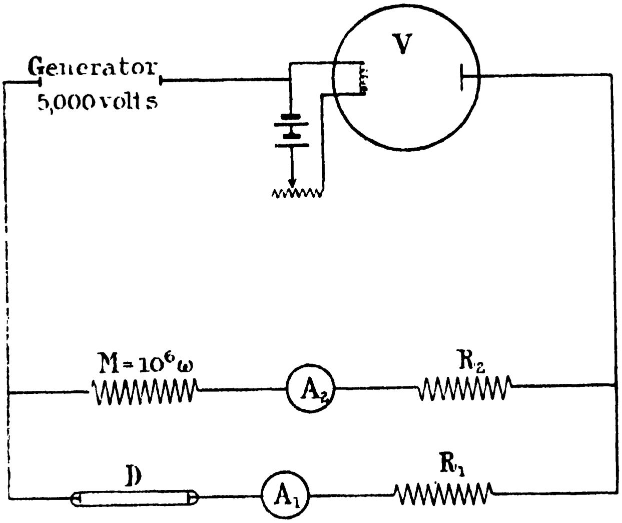


Fig. 1.

Compared to the resistance of the discharge tube, which was always greater than 1 megohm, and that of the megohm, the resistances (R_1+A_1) and (R_2+A_2) could always be neglected, since neither was greater than 300 ohms.

The source of power was an Evershed 5000 volt 15 milliampère direct current generator. The variations in the voltage of this generator were a serious source of trouble until a valve V was inserted in the circuit. I am indebted to Dr. Langmuir who suggested this important improvement. The current through the filament of this valve was so arranged that under the conditions of working the total current through the megohm and discharge tube was the saturation current of the valve. Any variations in the voltage of the generator including those small ones due to imperfect rectification were thus eliminated. The remaining cause of unsteadiness was the variation of the conditions prevailing in the discharge tube, and this was made as small as possible by the gas circulating device already described.

Insulation.

In order to ensure that insulation troubles should not arise, all the leads were suspended from glass tubes or were attached to blocks of paraffin wax or ebonite. The leads into the flask itself passed through separate glass tubes through the wooden top.

The insulation between the discharge tube circuit and the thermocouple circuit was tested by closing the latter circuit and switching on the generator. No deflection of the galvanometer could be observed.

Potentiometer.

The electromotive force developed by the thermocouples was measured on a Tinsley vernier potentiometer. This instrument was also used to measure the potential differences required for calibrations. It was not calibrated, as it had not been used since the makers had calibrated it and found it correct to at least 1 part in 10,000. A suspended coil galvanometer of the Ayrton-Mather type was used as an indicating instrument. The standard of E.M.F. consisted of two Weston cells by Tinsley.

Experimental Method.

The flask contained a heating coil of manganin wire. A measurement was made of the energy that had to be supplied to this coil to produce the same rise of temperature of the oil as was produced by the discharge tube working under known current and voltage. The electrical energies supplied in the two cases were compared. Since the thermal capacity was the same these two amounts of energy should agree, unless in the case of the hydrogen discharge changes occur which give rise to an evolution of energy. This comparative method made calibration of the thermocouples unnecessary. It is equivalent to determining the thermal capacity of the flask and its contents by means of the wire heating coil.

Measurement of Supply of Electrical Energy.

The energy supplied to the discharge tube was measured by reading the milliammeters A_1 and A_2 (fig. 1) as frequently as possible during the course of an experiment. These readings determine respectively the current through the tube and that through the megohm in parallel with the tube which gives the voltage drop over the tube. The product of the two readings was taken and the mean of two successive values of this product multiplied by the time interval between them. By adding together all the numbers so obtained we get the integrated energy supply for the whole experiment.

In the case of the wire heater, H (fig. 2), the resistance of the heater was determined, and the current through it measured by finding on the potentiometer the potential drop over the resistance R_3.

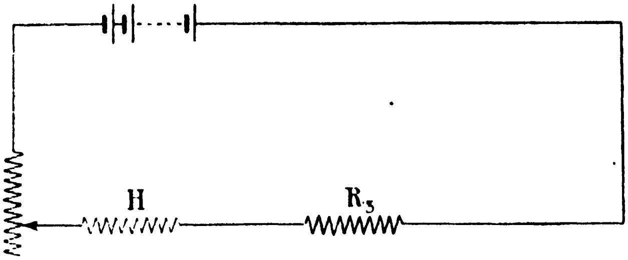


Fig. 2.

The resistances R_1 and R_2 which were of manganin and were oil immersed were used to calibrate A_1 and A_2 respectively. This calibration was effected by short-circuiting the valve V, replacing the generator by a single cell and a variable resistance, and short-circuiting the discharge tube to calibrate A_1 and the megohm to calibrate A_2. The instruments were read and the corresponding potential drops over R_1 and R_2 determined on the potentiometer.

To compare the results of heating by the wire and the tube it is necessary to know R_1, R_2, R_3, H and the megohm M in terms of one another. All these comparisons were made using the potentiometer.

$$H = \left.\begin{matrix} 10{\cdot}765 \\ 10{\cdot}764 \end{matrix}\right\} \text{mean} = 10{\cdot}76 \text{ ohms},$$

$$R_3 = \left.\begin{matrix} 0{\cdot}80133 \\ 0{\cdot}80133 \end{matrix}\right\} \text{mean} = 0{\cdot}80133 \text{ ohms}.$$

These were both compared directly with a standard 1 ohm coil by Tinsley.

$$R_1 = \left.\begin{matrix} 104{\cdot}66 \\ 104{\cdot}66 \end{matrix}\right\} \text{mean} = 104{\cdot}66 \text{ ohms},$$

$$R_2 = \left.\begin{matrix} 212{\cdot}04 \\ 212{\cdot}04 \end{matrix}\right\} \text{mean} = 212{\cdot}04 \text{ ohms}.$$

These were each determined by comparison with a 10 ohm manganin coil which was directly compared with the standard 1 ohm. Their ratio was checked at intervals. The megohm was subdivided into 20 coils of

50,000 ohms. Each of these coils was compared directly with R_2 with an accuracy of 1 part in 5000, so that if all the errors were in the same direction the maximum possible error in the sum would be 1 part in 250. The values obtained were :—

Coil.	Resistance.	Coil.	Resistance.
1	50,030	11	50,000
2	49,990	12	50,010
3	50,000	13	49,990
4	49,980	14	50,010
5	49,990	15	50,010
6	50,010	16	50,010
7	50,030	17	50,060
8	50,030	18	50,000
9	49,990	19	50,000
10	50,040	20	50,000

Total M = 1,000,180 ohms.

Carrying Capacity of Heating Coil and Megohm.

It was necessary to determine whether, under the conditions of the experiment, there was any appreciable change in the resistance of the heating coil, H, and the megohm, M, due to the heating effect of the currents passing through them. In the case of the heating coil this was tested by making a coil of the same gauge manganin wire as the coil itself was made of whose resistance was nearly the same as that of another coil of much thicker wire. The two coils were connected in series and immersed in oil. The ratio of the potential drops over them was determined on the potentiometer for different currents flowing through them. It was found when the current was varied from 0·2 to 0·72 ampères, that the ratio of the resistances did not change by 1 part in 1000. This shows that over this range of currents the change in resistance of the heating coil is inappreciable. This is practically the same as the range of currents used in the experiments, which was 0·1 to 0·8 ampères.

To test the megohm it was arranged so that one of the twenty coils was in the arm AD of the Wheatstone's bridge, fig. 3. The other nineteen coils were in the arm AB. The arms DC and BC were 100 and 1900 ohms respectively, and were of much heavier wire than the megohm. The change in the ratio of the resistances of AB and AD was less than 1 part in 1000 when the potential applied to the bridge

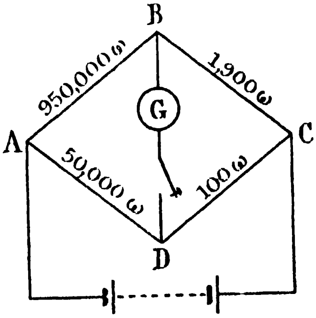


Fig. 3.

changed from 20 to 200 volts. This means that the resistance of AD did not change appreciably when the current through it was varied from 0·4 to 4 milliampères, since the change of current in the arm AB was inappreciable compared to this change. The maximum current passing through the megohm in the experiments performed was 3 milliampères.

Calculation of Energy Supply.

The method of determining the energy supply will be illustrated by giving, in tabular form, the calculations for a typical experiment. The second column contains the product of the observed deflections of the ammeters A_1 and A_2. The calibration factors are applied later.

Time.	Product of readings of A_1 and A_2.	Interval (T).	Mean product during interval (P).	P × T.
′ ″		″		
12 22	238·8			
14 37	236·4	(135)42	237·6	0·998 × 10^4
16 20	232·4	103	234·4	2·414
18 50	227·3	150	229·9	3·448
20 37	224·7	107	226·0	2·418
22 40	222·4	123	223·5	2·749
24 45	219·4	125	220·9	2·761
27 10	216·1	145	217·8	3·159
29 17	213·3	127	214·7	2·727
31 47	208·3	150	210·8	3·161
33 37	206·3	110	207·3	2·280
36 35	201·5	178	203·9	3·629
38 20	199·4	105	200·5	2·105
40 47	194·4	(147)99	196·9	1·949

Sum = 33·798 × 10^4,
= 3·380 × 10^5.

It will be noticed that in the cases of the first and last interval only part of the interval has been used as multiplier. This is because the rise of temperature was measured from 13 minutes 55 seconds to 39 minutes 59 seconds.

To reduce this result to watt-seconds we used the calibration factors of A_1, A_2. If e_1 and e_2 respectively are the voltage drops, as measured on the potentiometer over R_1 and R_2, corresponding to unit deflection of A_1 and A_2, we have the energy supply in watt seconds

$$= \frac{3{\cdot}380 \times 10^5 \times e_1 \times e_2}{R_1 \times R_2} \times 10^6.$$

It was found that for the shunts used in this experiment

$$e_1 = 0{\cdot}02811\,; \qquad e_2 = 0{\cdot}01711.$$

Inserting these values, and the values of R_1 and R_2, namely, 104·7 and 212·0, we obtain

$$\text{Energy supply} = \frac{3{\cdot}380 \times 10^5 \times 0{\cdot}02811 \times 0{\cdot}01711}{104{\cdot}7 \times 212{\cdot}0} \times 10^6$$

$$= 7{\cdot}324 \times 10^3 \text{ watt-seconds.}$$

Correction for Losses of Heat.

Changes in the temperature of the oil in the calorimeter may be caused by four factors other than the heat generated by the current in the discharge tube:—

(1) Communication of heat between the oil in the flask and the thermostat.

(2) Conduction through the top of the flask to the air.

(3) Heat generated by stirring.

(4) The temperature of the discharge tube will be above that of the oil in the flask, because heat is flowing from it to the oil. The excess temperature of the tube will depend on the power supply, since this determines the rate of flow of heat from the tube to the oil. If the power supply changes the excess temperature of the discharge tube will change. In other words, the amount of heat energy stored in the tube will change. This will be equivalent to a loss or gain of heat.

The corrections for the first three losses can be considered together as they were determined by the same method. The rate of change of thermocouple reading, with no current through the discharge tube, was plotted against thermocouple reading. This curve, which was a straight line, was drawn for the following standard conditions:—

Reading of Beckmann thermometer ... = 3·35 (22·7° C.)
Air temperature .... = 16°
Rate of stirring .... = 4 revolutions per second.

If all the experiments had been done under these standard conditions, the change in thermocouple reading due to losses during the course of an experiment could have been obtained from the graph. Actually, corrections had to be made to the rates of change determined from the graph for variations from the standard conditions.

The temperature of the thermostat was not the same throughout the whole series of experiments, owing to the fact that the regulator had to be readjusted several times. Correction was made for these changes by determining the change in thermocouple E.M.F. corresponding to 1/100°. This was found to be $0{\cdot}00206 \times 10^{-3}$ volts per 1/100°. Then for a thermostat temperature above 3·35, for example, 3·43 or 8/100° above, we subtract $8 \times 0{\cdot}00206 \times 10^{-3}$ from all the thermocouple readings, and read the rates of changes from the graph for the readings so obtained. Since the rate of change depends only on the difference between the oil and thermostat temperature, the values of the rate of change so obtained will be corrected for variations in the thermostat temperature from the standard. In other words, raising the thermostat temperature produces the same effect as lowering the oil temperature by the same amount.

Curves of rate of change of thermocouple reading plotted against thermocouple reading were drawn for different rates of stirring and for different air temperatures. By comparing curves in which (*a*) the variation of rate of stirring was large and that of air temperature small, (*b*) the variation of air temperature was large and that of rate of stirring small, it was possible to separate the two effects. The two groups of experiments will be considered separately :—

(*a*) In the experiments to determine the effect of changes in the rate of stirring on the rate of change of thermocouple reading, slight variations of the air temperature occurred. To correct for these variations, an approximate value for the effect of change of air temperature was required. This was obtained from two sets of readings at 22·05° C. and 16°–17° C. The approximate value was that 1° change in the air temperature produces a change of $0{\cdot}0000048 \times 10^{-3}$ volts per second in the rate of change of thermocouple readings. This was almost identical with the final corrected value 0·0000049. Applying this correction, the following results were obtained :—

Change in rate of stirring. δs.	Corresponding change in rate of change of thermo-couple reading. Δ.	$\Delta/\delta s$.
1·37 revs. per sec.	$0{\cdot}0000062 \times 10^{-3}$ volts per sec.	0·0000045
1·70 ,, ,,	0·0000075 ,, ,,	0·0000044
2·95 ,, ,,	0·0000096 ,, ,,	0·0000033
2·64 ,, ,,	0·0000157 ,, ,,	0·0000059

The numbers in the last column show that the correction to the rate of change for variations in the stirring may be taken as proportional to the variation in the rate of stirring. The mean of the numbers in the last column is 0·0000045. The error in this mean is certainly less than 0·000001. The two largest corrections to the rise in temperature of the oil due to variation in the rate of stirring from 4 revolutions per second amounted to 3 and 2·4 per cent. of the total rise. An error of 0·000001 in this figure would produce errors in the final results of 0·7 and 0·5 per cent. respectively. The actual errors due to the uncertainty of this correction will therefore be less than these.

(*b*) In the experiments to determine the effect of changes in the air temperature variations in the rate of stirring were corrected for among the figures just obtained. It was found that a change of 1° in the air temperature produced a change in the rate of change of thermocouple reading of $0{\cdot}0000049 \times 10^{-3}$ volts per second. This was the mean of twenty-three observations. The average departure from the mean was 0·0000005. The two largest corrections to the rise in temperature of the oil due to variations of the air temperature from 16° C. amount to 3 per cent., and a little less than 3 per cent. of the total rise. An error of 0·0000005 in the figure used to calculate these corrections would produce errors in the final result of 0·3 per cent. The uncertainty of this correction is less than this mean departure of the individual values.

Using these two correcting factors, all the observed rates of change were corrected to what they would have been if the rate of stirring had been 4 revolutions per second and the air temperature 16° C. The values so obtained were plotted against thermocouple readings, corrected, as explained above, for variations in the temperature of the thermostat from 3·35. This gave the heat-loss curve for the standard conditions. The points were found to lie on a straight line over the range of thermocouple readings $3{\cdot}4 \times 10^{-3}$ volts to $5{\cdot}9 \times 10^{-3}$ volts, that is, from 17° to 29° C.

We now consider the correction for changes in the amount of heat stored in the discharge tube due to changes in the power supply. This was determined by finding the amount of heat liberated when the power supply was changed by a known amount. To do this the following experiments were carried out: The scale of the galvanometer was calibrated. The discharge tube was run, the power supply being 6·86 watts. The current was shut off at a known instant, and the reading of the galvanometer scale noted at this time. Readings of the scale were taken at short intervals after this, and it was found that the oil temperature continued to rise after the power had been shut off. The deflections were reduced to thermocouple E.M.F.'s, and the values corrected for losses in the interval with the following results:—

Time in seconds from shutting off power.	Thermocouple E.M.F. (volts).
0	$5{\cdot}1592 \times 10^{-3}$
10	5·1652
25	5·1725
35	5·1745
50	5·1769
65	5·1785
80	5·1793
90	5·1798
100	5·1814

We see that when the power supply decreases by 6·86 watts the heat energy which is stored in the tube decreases, and that it causes a rise of $0·022 \times 10^{-3}$ volts in the thermocouple E.M.F. when it is distributed.

The correction may be taken as proportional to the change in power supply and to be equal to $0·00031 \times 10^{-3}$ volts for a change of 1 watt.

In applying this correction to the measured rise of temperature it is necessary to consider only the difference between the initial and final values of the power supplied.

Calculation of Heat-Loss Corrections.

The application of these corrections will be shown by taking a typical experiment and giving the complete reduction of the results.

Adding the figures in the last column (see Table facing) we obtain the total change of thermocouple E.M.F. due to the passage of heat through the walls and top of the flask $= -0·0455 \times 10^{-3}$ volts.

We now correct for stirring, remembering that the graph has been drawn for stirring at the rate of four revolutions per second :—

Interval.	Mean rate of revolution of stirrer in interval.	Rate of change of thermo. E.M.F. due to difference of rate of stirring from 4 revs. per sec.	Change of thermo. E.M.F. in interval due to excess stirring.
257	4·50	+0·0000023	+0·0006
206	4·70	32	+0·0007
287	5·01	45	+0·0012
292	4·27	12	+0·0003
290	4·45	20	+0·0006
252	5·06	48	+0·0012

Adding the figures in the last column we obtain the total change of thermocouple E.M.F. due to excess rate of stirring $= +0·0046 \times 10^{-3}$ volts.

The power at the end of the experiment was less than that at the beginning by 0·95 watts. We saw that a decrease of 1 watt produces an increase in the temperature of the oil as measured by the thermocouple E.M.F. of $0·0031 \times 10^{-3}$ volts. We have, therefore, the change of thermocouple E.M.F. due to change in the supply of power $= +0·0029 \times 10^{-3}$ volts.

Time.		Thermo-couple E.M.F. (volts × 10^{-4}).	Air temperature.	Thermostat temperature, Beckmann thermometer.	Interval T (seconds).	Mean thermo. E.M.F. during interval.	Mean thermo-couple E.M.F. corrected for variation of thermostat temperature from 3·35.	Rate of change of thermocouple from graph drawn for standard conditions.	Correction to rate of change for variation of air temperature from 16° C.	Rate of change corrected for variation of air temperature.	Change due to loss to thermostat and air in interval T (volts × 10^{-6}).
′	″										
13	55	4·2462	16·55	3·395							
					257	4·3113	4·3007	+0·0000140	+0·0000022	+0·0000162	+0·0041
18	12	4·3764	16·35	3·39							
					206	4·4265	4·4182	−0·0000038	18	−0·0000020	−0·0004
21	38	4·4765	16·40	3·39							
					267	4·5366	4·5284	−0·0000208	19	−0·0000189	−0·0050
26	5	4·5967	16·40	3·39							
					292	4·6568	4·6486	−0·0000395	24	−0·0000371	−0·0108
30	57	4·7169	16·60	3·39							
					290	4·7720	4·7638	−0·0000571	25	−0·0000546	−0·0158
35	47	4·8270	16·45	3·39							
					252	4·8721	4·8639	−0·0000725	25	−0·0000700	−0·0176
39	59	4·9171	16·60	3·39							

Collecting these results we have

Change of thermo. E.M.F. from 13′ 55″ to 39′ 59″ $= 4{\cdot}9171 - 4{\cdot}2462$	$= +0{\cdot}6709 \times 10^{-3}$ volts.
Change due to passage of heat through walls and top of flask	$= -0{\cdot}0455 \times 10^{-3}$,,
Change due to excess stirring above 4 revolutions per second	$= +0{\cdot}0046 + 10^{-3}$,,
Change due to decrease of power supply	$= +0{\cdot}0029 \times 10^{-3}$,,

We therefore have for $\delta\theta$ the change of thermocouple E.M.F. that would have been produced by the heat generated in the discharge tube alone.

$$\delta\theta = (0{\cdot}6709 + 0{\cdot}0455 - 0{\cdot}0046 - 0{\cdot}0029) \times 10^{-3} \text{ volts.}$$
$$= 0{\cdot}7089 \times 10^{-3} \text{ volts.}$$

Summary of Results.

In the following Tables the results which were obtained are summarised. The first Table refers to the experiments with the hydrogen discharge tube. The numbers in the sixth column, headed $\delta\theta$, give the change of thermocouple E.M.F. corrected for all losses. Those in the second last column give the total energy supplied to produce this change. In the last column is given the energy required to produce a change of 10^{-3} volts in the thermocouple E.M.F. In deducing this figure we assume that the relation between thermo. E.M.F. and temperature is approximately linear over the range of temperatures used. This range was 3° to 5°. We also assume that the specific heat of the oil is constant.

It will be noticed that there is no tendency for the figures in the last column to vary in a systematic way with the voltage. Their mean is $1{\cdot}035 \times 10^{-4}$ watt-seconds. The average departure from the mean is $0{\cdot}004 \times 10^{-4}$. It should be mentioned that there is no tendency for the results to vary in a systematic way with the time of the experiment, the air temperature, or the stirring. This shows that the corrections made for losses are sufficiently accurate.

In the case of the experiments in which the oil was heated by the wire heater, the following results were obtained for the energy required to produce a change of 10^{-3} volts in the thermocouple reading:—

Watt-seconds.	Watt-seconds.
$1{\cdot}043 \times 10^{4}$	$1{\cdot}040 \times 10^{4}$
$1{\cdot}039 \times 10^{4}$	$1{\cdot}041 \times 10^{4}$
$1{\cdot}034 \times 10^{4}$	$1{\cdot}029 \times 10^{4}$

Hydrogen Experiments.

Time.	Measured change of thermo. E.M.F. (volts × 10^{-3}).	Change due to		Decrease of power supply (volts × 10^{-3}).	$\delta\theta$ (volts × 10^{-3}).	Mean voltage.	Mean current (× 10^{-3} ampères).	Energy supply (watt-secs.).	Energy supply.
		Conduction of heat (volts × 10^{-3}).	Excess stirring (volts × 10^{-3}).						Corrected change of temperature.
″ 2675	0·6916	−0·0496	+0·0012	+0·0025	0·7375	866	3·34	7·696 × 10^3	1·043 × 10^{-4}
2185	0·6506	−0·0166	−0·0004	+0·0028	0·6648	979	3·18	6·873	1·035
1856	0·7391	−0·0509	+0·0110	+0·0041	0·7749	1150	3·71	8·063	1·040
1564	0·6709	−0·0455	+0·0046	+0·0029	0·7089	1363	3·45	7·324	1·033
1293	0·8526	−0·0248	+0·0064	+0·0028	0·8682	1617	4·30	8·991	1·035
1511	0·8600	−0·0259	+0·0110	+0·0034	0·8715	1652	3·61	8·999	1·033
1122	0·8635	−0·0224	+0·0064	+0·0029	0·8766	1670	4·39	9·053	1·032
1644	0·8892	−0·0258	+0·0125	+0·0137	0·8987	1713	3·30	9·272	1·033
2420	0·7628	−0·0379	+0·0177	—	0·7830	2050	1·64	8·058	1·029
2170	0·7712	−0·0660	+0·0255	+0·0009	0·8108	2260	1·72	8·356	1·030
1473	0·6203	−0·0275	−0·0022	—	0·6510	2716	1·70	6·798	1·044

These experiments were done under a variety of conditions as regards duration of experiment and air temperature. The results again did not show any systematic variation. Their mean is $1{\cdot}038 \times 10^4$ watt-seconds and the average departure is 0·004 watt-seconds.

It will be seen that the energy supply to produce the same rise of temperature is the same in the two cases to within the error of the experiment, the two values being $1{\cdot}035 \times 10^4$ watt seconds for the hydrogen discharge tube and $1{\cdot}038 \times 10^4$ watt-seconds for the wire heater. These values agreed with the value obtained in two or three experiments in which the tube was filled with air. This value was $1{\cdot}036 \times 10^4$ watt-seconds.

Discussion of Results.

It is interesting to note that these experiments prove that the conservation of energy holds in the energy changes in a discharge tube.

It would probably be possible to increase the accuracy of the heat measurements to 1 part in 500. This, however, would be extremely difficult, as will be seen by considering the very unfavourable conditions as regards quantity of heat per unit volume and remembering that under the most favourable conditions calorimetric measurements seldom attain an accuracy greater than 1 part in 1000. It would not be worth while doing this until a perfectly steady discharge could be obtained. At present this appears to be impossible owing to the large number of variables involved. In the present work two experiments had to be rejected on account of this unsteadiness. In many other cases conditions were so bad that it was impossible to obtain any measurements at all.

In conclusion it is a pleasure to thank Sir Ernest Rutherford for suggesting the experiment and for the interest he has taken in its progress. My thanks are also due to Mr. E. A. Milne of Trinity College for supplying me with the astrophysical data.

Free Motion of a Sphere in a Rotating Liquid parallel to the Axis of Rotation.

By S. F. Grace, University of Liverpool.

(Communicated by G. I. Taylor, F.R.S. Received May 1, 1922.)

§ 1. The motion is supposed to be a small disturbance from one of uniform rotation like a rigid body, small motion being defined to be such that the squares and products of velocity and vorticity components may be neglected in the expressions for acceleration. The system is supposed free from bodily forces, and is initially disturbed from relative rest by a motion suddenly communicated to the sphere.

The pressure intensity of the liquid consists of two parts, one depending only on the distance from the axis, and the other on the disturbed motion. If at any instant the sphere is moving parallel to the axis of rotation, and the disturbed motion of the liquid is symmetrical with respect to a line through the centre of the sphere parallel to the axis of rotation, the motional part of the pressure intensity will also be symmetrical with respect to this line, and its resultant effect on the sphere will be to produce an acceleration parallel to the axis of rotation. Now the positional part of the pressure intensity would maintain any portion of the liquid in relative rest, and will consequently maintain the sphere in relative rest, providing the density of the sphere is equal to that of the liquid, which we shall suppose to be the case.

The initial disturbed motion of the liquid will be irrotational, since the effects of rotation take time to develop, and if the initial relative motion of the sphere be parallel to the axis of rotation, the initial disturbed motion of the liquid will have the symmetrical character above specified.

It follows that the sphere will continue to move parallel to the axis of rotation, and the disturbed motion of the liquid will continue to be symmetrical with respect to the line of the sphere's relative motion, which in the sequel is called the axis of the sphere.

The method of discussion will be to assume the particular type of motion indicated above, and then to show that all the conditions can be satisfied by making it a particular function of the time. This leads to a perfectly definite mathematical problem, the solution of which gives the following results:—

The sphere oscillates about a point on its axis, the distance of which, from the initial position of the centre of the sphere, is proportional to the velocity of projection; the amplitude of the oscillation tends rapidly to zero, but the period tends to a constant value which is half that of the rotation of the undisturbed liquid.

With regard to the velocity of the liquid a general expression has been obtained for its value at any point, reducing to simple expressions in particular cases. These have been examined.

On the axis of the sphere the velocity of the liquid is always less than the initial velocity of projection of the sphere and tends to zero at infinity. In the neighbourhood of the sphere the velocity oscillates about the zero value with a period which is ultimately constant and equal to that of the sphere's oscillation; the amplitude also tends to a value which is constant at a particular point, but there is an increasingly large gradient of velocity along the axis.

On the equatorial plane of the sphere it appears that the disturbance is practically confined to the sphere's immediate neighbourhood. Away from the sphere the velocity of the liquid rapidly tends to zero, oscillating about the zero value with decreasing amplitude and a period which is ultimately constant at a particular point. At infinity this period becomes that of the ultimate oscillation of the sphere, but increases without limit as the sphere is approached. The velocity gradient over the plane also increases without limit.

Finally the velocity of the liquid tangential to the sphere rapidly approaches a value which is constant in magnitude at any point, varying from zero at the poles to a maximum at the equator, the direction, however, changes with constant angular velocity which has a maximum value at the poles and is zero at the equator. Here again there is an unlimited increase in velocity gradient, in this case along a meridian of the sphere.

However, since the gradients of the velocity of the liquid over the equatorial plane and of the transverse velocity along a meridian of the sphere ultimately increase without limit, parts of this solution are not applicable, for a stage is reached after which the components of vorticity cannot be considered as small. Consequently the physical assumptions underlying the mathematical treatment, viz., those of small motion, are violated; in other words, the solution will only represent the true state of the liquid for a restricted time.

The question of the ultimate physical state thus remains unanswered. J. Proudman* showed that a small steady disturbance was impossible, but then G. I. Taylor† obtained a solution of the general equations of steady symmetrical motion about a sphere, which does not satisfy the equations of small motion. Whether or not Taylor's solution represents the ultimate state in the problem of the present paper is a question for further investigation.

---

* "On the Motion of Solids in a Liquid Possessing Vorticity," 'Roy. Soc. Proc.,' A, vol. 92, 1916.

† 'Proc. Camb. Phil. Soc.,' vol. 20, Part III.

The work was sugggested to me by Prof. J. Proudman, to whom I am greatly indebted for his interest and advice; it is an extension of some of his own work (*loc. cit.*). It is proposed in a future paper to give the corresponding results when the sphere is projected at right angles to the axis of rotation, part of the initial work for which has already been effected. A combination of the two cases will give the small general motion of a sphere projected with given initial velocity.

Mathematical Problem.

§ 2. We shall use Cartesian axes $Oxyz$ rotating about Oz with constant angular velocity ω, called the primary axes. Let u, v, w denote the components of the velocity at x, y, z relative to these axes. The equation of continuity for a uniform liquid gives

$$\frac{\partial u}{\partial x}+\frac{\partial v}{\partial y}+\frac{\partial w}{\partial z}=0 \tag{1}$$

and the boundary condition may be written

$$lu+mv+nw=n\mathrm{V}, \tag{2}$$

l, m, n being the direction cosines of the normal to the sphere and V the velocity of the sphere. These relations being kinematical are independent of the fact that the axes are rotating.

The dynamical equations, in the case of small motion under no forces other than the pressure of the liquid, reduce to

$$\left.\begin{aligned}\frac{\partial u}{\partial t}-2\omega v &= -\frac{\partial \mathrm{P}}{\partial x}\\ \frac{\partial v}{\partial t}+2\omega u &= -\frac{\partial \mathrm{P}}{\partial y}\\ \frac{\partial w}{\partial t} &= -\frac{\partial \mathrm{P}}{\partial z}\end{aligned}\right\}, \tag{3}$$

where $$\mathrm{P}=p/\rho-\tfrac{1}{2}\omega^2(x^2+y^2), \tag{4}$$

the letters p and ρ having their usual signification.

Take a set of axes, having their origin always at the centre of the sphere (the secondary axes), and let these secondary axes be always parallel to the primary axes.

In equations (1)-(4) the actual values of x, y, z are only required in the term $\frac{1}{2}\omega^2(x^2+y^2)$ in (4), since in $\partial/\partial x$, $\partial/\partial y$, $\partial/\partial z$ we have t constant. Hence if x, y, z now refer to the secondary axes, the equations (1), (2), (3) remain, but

$$P = p/\rho - \tfrac{1}{2}\omega^2(x_0^2+y_0^2),$$

x_0, y_0, z_0, referring to the primary axes.

The term $\frac{1}{2}\rho\omega^2(x_0^2+y_0^2)$ in

$$p = \rho P + \tfrac{1}{2}\rho\omega^2(x_0^2+y_0^2)$$

would maintain a liquid sphere in relative rest, *i.e.*, would produce an acceleration $\omega^2(x_0^2+y_0^2)^{\frac{1}{2}}$ towards the axis of rotation. It will therefore produce an equal acceleration in the solid sphere, and we may take

$$p = \rho P$$

as the effective pressure intensity in producing the acceleration of the sphere other than centripetal.

Now assume that the functions u, v, w, V, and P may be expressed as power series in time. This proves to be legitimate, since the results obtained show these series to be highly convergent for finite values of t. Let

$$u = \sum_{s=0}^{\infty} u_s \frac{(2\omega t)^s}{s\,!}, \qquad v = \sum_{s=0}^{\infty} v_s \frac{(2\omega t)^s}{s\,!}, \qquad w = \sum_{s=0}^{\infty} w_s \frac{(2\omega t)^s}{s\,!},$$

$$V = \sum_{s=0}^{\infty} c_s \frac{(2\omega t)^s}{s\,!}, \qquad P = 2\omega \sum_{s=0}^{\infty} P_s \frac{(2\omega t)^s}{s\,!},$$

where u_s, v_s, w_s, c_s, and P_s are functions of x, y, z only, the initial state being given by u_0, v_0, w_0, c_0.

Substituting these expressions in the equations (1), (2), (3), we obtain

$$\frac{\partial u_s}{\partial x}+\frac{\partial v_s}{\partial y}+\frac{\partial w_s}{\partial z} = 0, \qquad s = 0, 1, 2, \ldots, \tag{5}$$

since the equation (1) is true for all time; the boundary condition gives

$$lu_s + mv_s + nw_s = nc_s, \qquad s = 0, 1, 2, \ldots,$$

and from the dynamical equations

$$\left.\begin{aligned} u_{s+1} - v_s &= -\partial P_s/\partial x \\ v_{s+1} + u_s &= -\partial P_s/\partial y \\ w_{s+1} \quad\; &= -\partial P_s/\partial z \end{aligned}\right\}, \qquad s = 0, 1, 2, \ldots, \tag{6}$$

from which

$$\left.\begin{aligned} u_{s+1} + u_{s-1} &= -\frac{\partial P_s}{\partial x} - \frac{\partial P_{s-1}}{\partial y} \\ v_{s+1} + v_{s-1} &= -\frac{\partial P_s}{\partial y} + \frac{\partial P_{s-1}}{\partial x} \\ w_{s+1} + w_{s-1} &= -\frac{\partial P_s}{\partial z} - \frac{\partial P_{s-2}}{\partial z} \end{aligned}\right\}, \qquad s = 2, 3, 4, \ldots, \tag{7}$$

which remain valid for $s = 1$ if we replace $\partial P_{-1}/\partial z$ by $-w_0$.

The velocity of the liquid at any point may now be expressed in terms of the sequence P_s, thus*

$$u - u_0\gamma_0(2\omega t) - v_0\gamma_1(2\omega t)$$

$$= -\partial/\partial x \sum_{s=0}^{\infty} P_s\gamma_{s+1}(2\omega t) - \partial/\partial y \sum_{s=0}^{\infty} P_s\gamma_{s+2}(2\omega t),$$

$$v - v_0\gamma_0(2\omega t) + u_0\gamma_1(2\omega t)$$

$$= -\partial/\partial y \sum_{s=0}^{\infty} P_s\gamma_{s+1}(2\omega t) + \partial/\partial x \sum_{s=0}^{\infty} P_s\gamma_{s+2}(2\omega t),$$

$$w - w_0 = -\partial/\partial z \sum_{s=0}^{\infty} P_s(2\omega t)^{s+1}/(s+1)!$$

where $$\gamma_s(z) = \sum_{n=0}^{\infty} (-1)^n z^{s+2n}/(s+2n)!$$

and we now concentrate our attention on the sequence P_s.

Inside the liquid we have, from (5) and (7),

$$\nabla^2 P_s = -\partial^2 P_{s-2}/\partial z^2, \quad s = 1, 2, 3 \ldots,$$

and from (6)

$$\nabla^2 P_0 = 0,$$

while at the boundary

$$\frac{\partial P_s}{\partial n} = -nc_{s+1} - nc_{s-1} - n\frac{\partial P_{s-2}}{\partial z}, \qquad s = 1, 2, 3, \ldots,$$

on account of the symmetrical character already referred to, with

$$\partial P_0/\partial n = -nc_1.$$

The initial motion being irrotational, we may let P_{-1} denote the velocity potential, so that

$$u_0 = -\partial P_{-1}/\partial x, \qquad v_0 = -\partial P_{-1}/\partial y, \qquad w_0 = -\partial P_{-1}/\partial z,$$

and $$\nabla^2 P_{-1} = 0, \qquad \partial P_{-1}/\partial n = -nc_0.$$

The value of c_0, the initial velocity of the sphere, determines the magnitude of the subsequent motion; in the sequel this quantity, c_0, is denoted by W.

The resultant pressure on the sphere is in the direction Oz and equal to

$$-\rho \iint (P)_{r=a} \cos\theta \, dS$$

taken over the sphere, where r, θ, ϕ are the spherical polar co-ordinates of the element dS referred to the centre of the sphere as origin, $\theta = 0$ denoting the z axis and $\phi = 0$ the plane xOz, a being the radius of the sphere. This becomes

$$-2\omega\rho \sum_{s=0}^{\infty} \frac{(2\omega t)^s}{s!} \iint (P_s)_{r=a} \cos\theta \, dS,$$

---

* Proudman, *loc. cit.*

and the equations of motion of the sphere reduce to

$$\sum_{s=0}^{\infty} \frac{(2\omega t)^s}{s!} \iint (P_s)_{r=a} \cos\theta \, dS = -\tfrac{4}{3}\pi a^3 \sum_{s=1}^{\infty} c_s \frac{(2\omega t)^{s-1}}{(s-1)!}$$

so that

$$c_{s+1} = -\iint (P_s)_{r=a} \cos\theta \, dS / (\tfrac{4}{3}\pi a^3), \qquad s = 0, 1, 2, \dots, \qquad (8)$$

giving the functions c_s in terms of the sequence P_s.

Assuming the liquid to be at relative rest at infinity, the boundary conditions may now be written.

$$\left(\frac{\partial P_s}{\partial r}\right)_{r=a} = \frac{\cos\theta}{\frac{4}{3}\pi a^3} \iint (P_s + P_{s-2})_{r=a} \cos\theta \, dS - \cos\theta \left(\frac{\partial P_{s-2}}{\partial z}\right)_{r=a}, \qquad s = 2, 3, \dots,$$

and

$$\left(\frac{\partial P_1}{\partial r}\right)_{r=a} = \frac{\cos\theta}{\frac{4}{3}\pi a^3} \iint (P_1)_{r=a} \cos\theta \, dS - W\cos\theta - \cos\theta \left(\frac{\partial P_{-1}}{\partial z}\right)_{r=a},$$

$$\left(\frac{\partial P_0}{\partial r}\right)_{r=a} = \frac{\cos\theta}{\frac{4}{3}\pi a^3} \iint (P_0)_{r=a} \cos\theta \, dS.$$

Now $\qquad \nabla^2 P_{-1} = 0, \qquad (\partial P_{-1}/\partial r)_{r=a} = -W\cos\theta,$

hence $\qquad P_{-1} = \tfrac{1}{2} W a (a/r)^2 \cos\theta$

and $\qquad \nabla^2 P_0 = 0, \qquad \left(\frac{\partial P_0}{\partial r}\right)_{r=a} = \frac{\cos\theta}{\frac{4}{3}\pi a^3} \iint (P_0)_{r=a} \cos\theta \, dS.$

The function P_0 has to satisfy these two relations and vanish at infinity; this is impossible unless P_0 is zero. Generally

$$\nabla^2 P_{2s} = 0, \qquad \left(\frac{\partial P_{2s}}{\partial r}\right)_{r=a} = \frac{\cos\theta}{\frac{4}{3}\pi a^3} \iint (P_{2s})_{r=a} \cos\theta \, dS,$$

$$\therefore \qquad P_{2s} = 0.$$

Hence we require a sequence of functions P_{2s-1} ($s = 1, 2, \dots$), which vanish at infinity and satisfy

$$\nabla^2 P_{2s-1} = -\partial^2 P_{2s-3}/\partial z^2, \qquad s = 2, 3, \dots, \qquad (9)$$

inside the liquid and

$$\left(\frac{\partial P_{2s-1}}{\partial r}\right)_{r=a} = \frac{\cos\theta}{\frac{4}{3}\pi a^3} \iint (P_{2s-1} + P_{2s-3})_{r=a} \cos\theta \, dS - \cos\theta \left(\frac{\partial P_{2s-3}}{\partial z}\right)_{r=a},$$

$$s = 2, 3, \dots, \qquad (10)$$

together with

$$\nabla^2 P_1 = -\partial^2/\partial z^2 \{ \tfrac{1}{2} W a (a/r)^2 \cos\theta \}, \qquad (11)$$

and

$$\left(\frac{\partial P_1}{\partial r}\right)_{r=a} = \frac{\cos\theta}{\frac{4}{3}\pi a^3} \iint (P_1)_{r=a} \cos\theta \, dS - W\cos\theta$$

$$- \cos\theta \left[\frac{\partial}{\partial z} \left\{ \tfrac{1}{2} W a \left(\frac{a}{r}\right)^2 \cos\theta \right\} \right]_{r=a}. \qquad (12)$$

The Solution.

§ 3. After an investigation which would take too much space to reproduce, the required sequence was found to be as follows:—

$$\mathrm{P}_{2s-1} = \mathrm{W}a\,\{(a/r)^2\,\mathrm{Q}_{2,\,2s}(\cos\theta)+(a/r)^4\,\mathrm{Q}_{4,\,2s}(\cos\theta)+\ldots$$

$$\ldots+(a/r)^{2m}\,\mathrm{Q}_{2m,\,2s}(\cos\theta)+\ldots+(a/r)^{2s+2}\,\mathrm{Q}_{2s+2,\,2s}(\cos\theta)\}, \quad (13)$$

where, putting $\cos\theta = \mu$,

$$\mathrm{Q}_{2m,\,2s}(\mu) = (-1)^{s-m}\,\frac{3}{2}\cdot\frac{1\,.\,3\,.\,5\,.\,\ldots\,(2s-1)}{2^s\,(m-1)\,!\,(s-m+1)\,!}\,\mu\,(1-\mu^2)^{s-m}\times$$

$$\left[\frac{2}{2m+1}\left\{1-\frac{2s+1}{3}\,{}^{m-1}\mathrm{C}_1\mu^2+\frac{2s+1}{3}\,.\,\frac{2s+3}{5}\,{}^{m-1}\mathrm{C}_2\mu^4-\ldots\right.\right.$$

$$\left.\ldots+(-1)^{m-1}\,\frac{2s+1}{3}\,.\,\frac{2s+3}{5}\ldots\frac{2s+2m-3}{2m-1}\,\mu^{2m-2}\right\}$$

$$-\left\{1-\frac{2s+1}{3}\,{}^{m}\mathrm{C}_1\mu^2+\frac{2s+1}{3}\,.\,\frac{2s+3}{5}\,{}^{m}\mathrm{C}_2\mu^4-\ldots\right.$$

$$\left.\left.\ldots+(-1)^{m}\,\frac{2s+1}{3}\,.\,\frac{2s+3}{5}\ldots\frac{2s+2m-1}{2m+1}\,\mu^{2m}\right\}\right], \quad (14)$$

for $s > 0$ and $m = 1, 2, \ldots (s+1)$. A more compact form is given by

$$\mathrm{Q}_{2m,\,2s}(\mu) = (-1)^{s-m}\,\frac{3}{4}\cdot\frac{(1-\mu^2)^{s-m}}{(m-1)\,!\,(s-m+1)\,!}\times$$

$$\left(\frac{d}{d\mu^2}\right)^{s-1}\left\{\frac{2\mu^{2s-1}}{2m+1}\,(1-\mu^2)^{m-1}-\mu^{2s-1}\,(1-\mu^2)^{m}\right\}, \quad (15)$$

and it will be noticed that

$$\mathrm{Q}_{2s+2,\,2s}(\mu) = (-1)^s\,\tfrac{3}{2}\,\mathrm{P}_{2s+1}(\mu)/(2s+3),$$

where, in this case, $\mathrm{P}_{2s+1}(\mu)$ denotes Legendre's function.

We shall now simply verify that these formulæ satisfy all the conditions.

Verification.

§ 4. Considering first the relation (9), the expression (13) satisfies this provided

$$\nabla^2\{(a/r)^{2m}\mathrm{Q}_{2m,\,2s}\} = -\partial^2/\partial z^2\{(a/r)^{2m}\mathrm{Q}_{2m,\,2s-2}\}, \qquad s = 2, 3, \ldots,$$

i.e.
$$(1-\mu^2)\,\frac{\partial^2\mathrm{Q}_{2m,\,2s}}{\partial\mu^2}-2\mu\,\frac{\partial\mathrm{Q}_{2m,\,2s}}{\partial\mu}+2m\,(2m-1)\,\mathrm{Q}_{2m,\,2s}$$

$$= -\left[(1-\mu^2)^2\,\frac{\partial^2\mathrm{Q}_{2m,\,2s-2}}{\partial\mu^2}-(4m+3)\,\mu\,(1-\mu^2)\,\frac{\partial\mathrm{Q}_{2m,\,2s-2}}{\partial\mu}\right.$$

$$\left.+2m\,\{2\,(m+1)\,\mu^2-1\}\,\mathrm{Q}_{2m,\,2s-2}\right].$$

This relation is satisfied when $m = s+1$ since $\mathrm{Q}_{2s+2,\,2s}$ is Legendre's function of order $(2s+1)$, apart from a factor, and $\mathrm{Q}_{2s+2,\,2s-2}$ is zero. Otherwise let

$$\mathrm{Q}_{2m,\,2s} = (-1)^{s-m}\,\frac{3}{2}\,\frac{1\,.\,3\,.\,5\,\ldots\,(2s-1)}{2^s\,(m-1)\,!\,(s-m+1)\,!}\,q_{2m,\,2s},$$

then we have to show that

$$(2s-1)\left[(1-\mu^2)\frac{\partial^2 q_{2m,2s}}{\partial\mu^2}-2\mu\frac{\partial q_{2m,2s}}{\partial\mu}+2m(2m-1)q_{2m,2s}\right]$$

$$=2(s-m+1)\left[(1-\mu)^2\frac{\partial^2 q_{2m,2s-2}}{\partial\mu^2}-(4m+3)\mu(1-\mu^2)\frac{\partial q_{2m,2s-2}}{\partial\mu}\right.$$

$$\left.+2m\left\{2(m+1)\mu^2-1\right\}q_{2m,2s-2}\right].$$

Now

$$q_{2m,2s}=(1-\mu^2)^{s-m}\left[\frac{2}{2m+1}\left\{\mu-\frac{2s+1}{3}{}^{m-1}C_1\mu^3+\frac{2s+1}{3}\cdot\frac{2s+3}{5}{}^{m-1}C_2\mu^5-\ldots\right.\right.$$

$$\ldots+(-1)^k\frac{2s+1}{3}\cdot\frac{2s+3}{5}\ldots\frac{2s+2k-1}{2k+1}{}^{m-1}C_k\mu^{2k+1}+\ldots$$

$$\left.\ldots+(-1)^{m-1}\frac{2s+1}{3}\cdot\frac{2s+3}{5}\ldots\frac{2s+2m-3}{2m-1}\mu^{2m-1}\right\}$$

$$\left.-\left\{\mu-\frac{2s+1}{3}{}^{m}C_1\mu^3+\ldots+(1)^m\frac{2s+1}{3}\ldots\frac{2s+2m-1}{2m+1}\mu^{2m+1}\right\}\right]$$

$$=(1-\mu^2)^{s-m}\alpha_{2s}\text{ say.}$$

$$\frac{\partial q_{2m,2s}}{\partial\mu}=-2(s-m)(1-\mu^2)^{s-m-1}\mu\alpha_{2s}$$

$$+(1-\mu^2)^{s-m}\left[\frac{2}{2m+1}\left\{1-\frac{2s+1}{3}{}^{m-1}C_1 3\mu^2\right.\right.$$

$$+\frac{2s+1}{3}\cdot\frac{2s+3}{5}{}^{m-1}C_2 5\mu^4-\ldots$$

$$\left.\ldots+(-1)^k\frac{2s+1}{3}\ldots\frac{2s+2k-1}{2k+1}{}^{m-1}C_k(2k+1)\mu^{2k}+\ldots\right\}$$

$$\left.-\left\{1-\frac{2s+1}{3}{}^{m}C_1 3\mu^2+\ldots\right\}\right]$$

$$=-2(s-m)(1-\mu^2)^{s-m-1}\mu\alpha_{2s}+(1-\mu^2)^{s-m}\beta_{2s}\text{ say.}$$

$$\frac{\partial^2 q_{2m,2s}}{\partial\mu^2}=4(s-m)(s-m-1)(1-\mu^2)^{s-m-2}\mu^2\alpha_{2s}$$

$$-2(s-m)(1-\mu^2)^{s-m-1}\alpha_{2s}-4(s-m)(1-\mu^2)^{s-m-1}\mu\beta_{2s}$$

$$+(1-\mu^2)^{s-m}\left[\frac{2}{2m+1}\left\{-\frac{2s+1}{3}{}^{m-1}C_1 3.2\mu\right.\right.$$

$$+\frac{2s+1}{3}\cdot\frac{2s+3}{5}{}^{m-1}C_2 5.4\mu^3-\ldots$$

$$\left.\ldots+(-1)^k\frac{2s+1}{3}\ldots\frac{2s+2k-1}{2k+1}{}^{m-1}C_k(2k+1)2k\mu^{2k-1}+\ldots\right\}$$

$$\left.-\left\{-\frac{2s+1}{3}{}^{m}C_1 3.2\mu+\ldots\right\}\right]$$

$$=2(s-m)\left\{(2s-2m-1)\mu^2-1\right\}(1-\mu^2)^{s-m-2}\alpha_{2s}$$

$$-4(s-m)(1-\mu^2)^{s-m-1}\mu\beta_{2s}+(1-\mu^2)^{s-m}\gamma_{2s}\text{ say.}$$

Hence

$$(1-\mu^2)\frac{\partial^2 q_{2m,\,2s}}{\partial\mu^2}-2\mu\frac{\partial q_{2m,\,2s}}{\partial\mu}+2m(2m-1)\,q_{2m,\,2s}$$

$$=(1-\mu^2)^{s-m-1}\left[\{2s(2s-4m+1)\mu^2+4m^2-2s\}\,\alpha_{2s}\right.$$
$$\left.-2(2s-2m+1)\mu(1-\mu^2)\,\beta_{2s}+(1-\mu^2)^2\gamma_{2s}\right]$$

and

$$(1-\mu^2)^2\frac{\partial^2 q_{2m,\,2s-2}}{\partial\mu^2}-(4m+3)\mu(1-\mu^2)\frac{\partial q_{2m,\,2s-2}}{\partial\mu}$$
$$+2m\{2(m+1)\mu^2-1\}\,q_{2m\,\,2s-2}$$

$$=(1-\mu^2)^{s-m-1}\left[\{4s(s-1)\mu^2-2(s-1)\}\,\alpha_{2s-2}\right.$$
$$\left.-(4s-1)\mu(1-\mu^2)\,\beta_{2s-2}+(1-\mu^2)^2\gamma_{2s-2}\right];$$

so that we are required to show that

$$(2s-1)\left[\{2s(2s-4m+1)\mu^2+4m^2-2s\}\,\alpha_{2s}\right.$$
$$\left.-2(2s-2m+1)\mu(1-\mu^2)\,\beta_{2s}+(1-\mu^2)^2\gamma_{2s}\right] \qquad (16)$$

is identical with

$$2(s-m+1)\left[\{4s(s-1)\mu^2-2(s-1)\}\,\alpha_{2s-2}\right.$$
$$\left.-(4s-1)\mu(1-\mu^2)\,\beta_{2s-2}+(1-\mu^2)^2\gamma_{2s-2}\right]. \qquad (17)$$

The coefficient of μ^{2k+1} in (16) is ($k = 2, 3, \ldots\, m-2$)

$$(-1)^{k+1}\left[\frac{2s-1}{1}\cdot\frac{2s+1}{3}\cdot\ldots\frac{2s+2k+1}{2k+3}\times\right.$$
$$\left(\frac{2}{2m+1}\,{}^{m-1}\mathrm{C}_{k+1}-{}^{m}\mathrm{C}_{k+1}\right)(2k+3)(2k+2)$$
$$+\frac{2s-1}{1}\cdot\ldots\frac{2s+2k-1}{2k+1}\left(\frac{2}{2m+1}\,{}^{m-1}\mathrm{C}_k-{}^{m}\mathrm{C}_k\right)\times$$
$$\{-4m^2+2s+2(2s-2m+1)(2k+1)+2(2k+1)2k\}$$
$$+\frac{2s-1}{1}\cdot\ldots\frac{2s+2k-3}{2k-1}\left(\frac{2}{2m+1}\,{}^{m-1}\mathrm{C}_{k-1}-{}^{m}\mathrm{C}_{k-1}\right)\times$$
$$\left.\{2s(2s-4m+1)+2(2s-2m+1)(2k-1)+(2k-1)(2k-2)\}\right]$$

$$=(-1)^{k+1}\frac{2s-1}{1}\cdot\frac{2s+1}{3}\cdot\ldots\frac{2s+2k-1}{2k+1}\times$$
$$\left\{4(s-m)\,{}^{m}\mathrm{C}_k-{}^{m}\mathrm{C}_{k-1}\,\frac{2(s-m+1)(2m+3)(m+1)-4(m-k+1)}{k}\right\}$$

$$=(-1)^{k+1}\frac{2s-1}{1}\cdot\frac{2s+1}{3}\cdot\ldots\frac{2s+2k-1}{2k+1}\,2(s-m+1)\times$$
$$\{2\,{}^{m}\mathrm{C}_k-(2m+3)\,{}^{m+1}\mathrm{C}_k\},$$

and the coefficient of μ^{2k+1} in (17) is

$$
\begin{aligned}
&(-1)^{k+1}\Big[2(s-m+1)\frac{2s-1}{3}\cdot\frac{2s+1}{5}\dots\frac{2s+2k-1}{2k+3}\times\\
&\qquad\qquad\left(\frac{2}{2m+1}\,{}^{m-1}C_{k+1}-{}^{m}C_{k+1}\right)(2k+3)(2k+2)\\
&\quad+2(s-m+1)\frac{2s-1}{3}\dots\frac{2s+2k-3}{2k+1}\left(\frac{2}{2m+1}\,{}^{m-1}C_{k}-{}^{m}C_{k}\right)\times\\
&\qquad\qquad\{2(s-1)+(4s-1)(2k+1)+2(2k+1)2k\}\\
&\quad+2(s-m+1)\frac{2s-1}{3}\dots\frac{2s+2k-5}{2k-1}\left(\frac{2}{2m+1}\,{}^{m-1}C_{k-1}-{}^{m}C_{k-1}\right)\times\\
&\qquad\qquad\{4s(s-1)+(4s-1)(2k-1)+(2k-1)(2k-2)\}\Big]\\
&=(-1)^{k+1}\,2(s-m+1)\frac{2s-1}{1}\cdot\frac{2s+1}{3}\cdot\dots\frac{2s+2k-1}{2k+1}\times\\
&\qquad\qquad\{2\,{}^{m}C_{k}-(2m+3)\,{}^{m+1}C_{k}\},
\end{aligned}
$$

and it is easily verified that the coefficients of $\mu, \mu^3, \mu^{2m+1}, \mu^{2m+3}$ in (16) and (17) are also identical.

Examining now the case where $s=1$, we have

$$
\begin{aligned}
\nabla^2 P_1 &= \nabla^2 Wa\left\{\tfrac{1}{4}(a/r)^2\mu(3\mu^2-1)-\tfrac{3}{20}(a/r)^4\mu(5\mu^2-3)\right\}\\
&= \tfrac{3}{2}W/a\,(a/r)^4\mu(3-5\mu^2)\\
&= -\partial^2/\partial z^2\left\{\tfrac{1}{2}Wa\,(a/r)^2\mu\right\}.
\end{aligned}
$$

It now remains to show that the assumed form (13) of P_{2s-1} satisfies the boundary conditions.

On $r=a$ we have

$$
P_{2s-1} = Wa\left\{Q_{2,2s}(\mu)+Q_{4,2s}(\mu)+\dots+Q_{2m,2s}(\mu)+\dots+Q_{2s+2,2s}(\mu)\right\},
$$

and, examining the series within the brackets, we have, putting $\mu^2=\nu$

$$
\begin{aligned}
&Q_{2,2s}+Q_{4,2s}+\dots+Q_{2m,2s}+\dots+Q_{2s+2,2s}\\
&=(-1)^{s-1}\frac{3}{4}\Big[\frac{(1-\nu)^{s-1}}{s\,!}\left(\frac{d}{d\nu}\right)^{s-1}\left\{\frac{2}{3}\,\nu^{(2s-1)/2}-\nu^{(2s-1)/2}(1-\nu)\right\}\\
&\quad-\frac{(1-\nu)^{s-2}}{1\,!\,(s-1)\,!}\left(\frac{d}{d\nu}\right)^{s-1}\left\{\frac{2}{5}\,\nu^{(2s-1)/2}(1-\nu)-\nu^{(2s-1)/2}(1-\nu)^2\right\}\\
&\quad+\frac{(1-\nu)^{s-3}}{2\,!\,(s-2)\,!}\left(\frac{d}{d\nu}\right)^{s-1}\left\{\frac{2}{7}\,\nu^{(2s-1)/2}(1-\nu)^2-\nu^{(2s-1)/2}(1-\nu)^3\right\}\\
&\quad\dots-(-1)^m\frac{(1-\nu)^{s-m}}{(m-1)\,!\,(s-m+1)\,!}\times\\
&\qquad\qquad\left(\frac{d}{d\nu}\right)^{s-1}\left\{\frac{2}{2m+1}\,\nu^{(2s-1)/2}(1-\nu)^{m-1}-\nu^{(2s-1)/2}(1-\nu)^m\right\}\\
&\quad\dots+(-1)^s\frac{(1-\nu)^{-1}}{s!}\left(\frac{d}{d\nu}\right)^{s-1}\left\{\frac{2}{2s+3}\,\nu^{(2s-1)/2}(1-\nu)^s-\nu^{(2s-1)/2}(1-\nu)^{s+1}\right\}\Big]
\end{aligned}
$$

$$= (-1)^{s-1} 3 \left[\frac{\nu^{1/2} (1-\nu)^{s-1}}{(2s+1)(2s+3)} + \frac{\nu^{3/2} (1-\nu)^{s-2}}{(2s+1)(2s+3)} {}^{s-1}C_1 + \dots \right]$$

$$= (-1)^{s-1} 3\mu / \{(2s+1)(2s+3)\}.$$

Hence

$$(P_{2s-1})_{r=a} = (-1)^{s-1} 3Wa \cos\theta / \{(2s+1)(2s+3)\}, \qquad (18)$$

and it follows from (8) that

$$c_{2s} = -\frac{1}{\frac{4}{3}\pi a^3} \iint (-1)^{s-1} \frac{3Wa}{(2s+1)(2s+3)} \cos^2\theta \, dS, \qquad s > 0$$

$$= (-1)^s 3W / \{(2s+1)(2s+3)\} \qquad (19)$$

and c_{2s-1} is zero. Next

$$(\partial P_{2s-1}/\partial r)_{r=a} = -W \{2Q_{2,2s} + 4Q_{4,2s} + \dots + 2m Q_{2m,2s} + \dots + (2s+2) Q_{2s+2,2s}\}$$

$$= (-1)^s \frac{3W}{4} \left[2 \frac{(1-\nu)^{s-1}}{s!} \left(\frac{d}{d\nu}\right)^{s-1} \left\{ \frac{2}{3} \nu^{(2s-1)/2} - \nu^{(2s-1)/2} (1-\nu) \right\} \right.$$

$$-4 \frac{(1-\nu)^{s-2}}{1!\,(s-1)!} \left(\frac{d}{d\nu}\right)^{s-1} \left\{ \frac{2}{5} \nu^{(2s-1)/2} (1-\nu) - \nu^{(2s-1)/2} (1-\nu)^2 \right\}$$

$$\dots -(-1)^m 2m \frac{(1-\nu)^{s-m}}{(m-1)!\,(s-m+1)!} \times$$

$$\left(\frac{d}{d\nu}\right)^{s-1} \left\{ \frac{2}{2m+1} \nu^{(2s-1)/2} (1-\nu)^{m-1} - \nu^{(2s-1)/2} (1-\nu)^m \right\}$$

$$\dots + (-1)^s (2s+2) \frac{(1-\nu)^{-1}}{s!} \times$$

$$\left. \left(\frac{d}{d\nu}\right)^{s-1} \left\{ \frac{2}{2s+3} \nu^{(2s-1)/2} (1-\nu)^s - \nu^{(2s-1)/2} (1-\nu)^{s+1} \right\} \right]$$

$$= (-1)^s \frac{3W}{4} \left[-\frac{4\nu^{\frac{1}{2}}}{(2s+1)(2s+3)} \times \right.$$

$$\left\{ (1-\nu)^{s-1} + {}^{s-1}C_1 \nu (1-\nu)^{s-2} + {}^{s-1}C_2 \nu^2 (1-\nu)^{s-3} + \right.$$

$$\left. \left. \dots + \nu^{s-1} \right\} + 2\nu^{(2s-1)/2} (1-\nu) \right]$$

$$= (-1)^s W \left[\tfrac{3}{2} \mu^{2s-1} (1-\mu^2) - 3\mu / \{(2s+1)(2s+3)\} \right], \qquad s > 0. \qquad (20)$$

Finally

$$\left(\frac{\partial P_{2s-1}}{\partial z}\right)_{r=a} = \mu \left(\frac{\partial P_{2s-1}}{\partial r}\right)_{r=a} + \frac{1-\mu^2}{a} \left(\frac{\partial P_{2s-1}}{\partial \mu}\right)_{r=a}$$

$$= (-1)^s W \left\{ \frac{3}{2} \mu^{2s} (1-\mu^2) - \frac{3\mu^2}{(2s+1)(2s+3)} - \frac{3(1-\mu^2)}{(2s+1)(2s+3)} \right\}$$

$$= (-1)^s W \left[\tfrac{3}{2} \mu^{2s} (1-\mu^2) - 3 / \{(2s+1)(2s+3)\} \right].$$

Substituting now in equation (10), we have

$$(-1)^s \mathrm{W}\left[\tfrac{3}{2}\mu^{2s-1}(1-\mu^2)-3\mu/\{(2s+1)(2s+3)\}\right]$$
$$= \mu\left\{(-1)^{s-1}\frac{3\mathrm{W}}{(2s+1)(2s+3)}+(-1)^{s-2}\frac{3\mathrm{W}}{(2s-1)(2s+1)}\right\}$$
$$-\mu(-1)^{s-1}\mathrm{W}\left\{\frac{3}{2}\mu^{2s-2}(1-\mu^2)-\frac{3}{(2s-1)(2s+1)}\right\}$$

for $s = 2, 3, \ldots$ and

$$(\partial \mathrm{P}_1/\partial r)_{r=a} = -\mathrm{W}\left\{\tfrac{3}{2}\mu(1-\mu^2)-\tfrac{3}{5}\mu\right\} = -\mathrm{W}\left\{\tfrac{13}{10}\mu-\tfrac{3}{2}\mu^3\right\}$$
$$= \frac{\cos\theta}{\frac{4}{3}\pi a^3}\iint(\mathrm{P}_1)_{r=a}\cos\theta\, d\mathrm{S}-\mathrm{W}\cos\theta-\cos\theta\left[\frac{\partial}{\partial z}\left\{\tfrac{1}{2}\mathrm{W}a\left(\frac{a}{r}\right)^2\cos\theta\right\}\right]_{r=a}.$$

It follows that the sequence P_{2s-1}, given in (13), satisfies the conditions of the problem, and therefore determines the motion.

Convergence.

§ 5. Before proceeding to the results obtained from this sequence, we shall consider the convergence of the series for u, v, w, and P, in terms of P_{2s-1}.

From the expression for $\mathrm{Q}_{2m, 2s}(\mu)$, given in (14), it is seen that, for $m \leq s$

$$|\mathrm{Q}_{2m, 2s}(\mu)| < \frac{3}{2}\cdot\frac{1}{2^s(m-1)!\,(s-m+1)!}\times$$
$$[2(2m+3)(2m+5)\ldots(2s+2m-3)\{1+{}^{m-1}\mathrm{C}_1+{}^{m-1}\mathrm{C}_2+\ldots\}$$
$$+(2m+3)(2m+5)\ldots(2s+2m-1)\{1+{}^{m}\mathrm{C}_1+{}^{m}\mathrm{C}_2+\ldots\}]$$
$$< \frac{3}{2}\cdot\frac{(2m+3)(2m+5)\ldots(2s+2m-3)}{2^{s-m}(m-1)!\,(s-m+1)!}(2s+2m)$$
$$< \frac{3}{2}2^{m-1}\frac{(s+m)(s+m-1)(s+m-2)\ldots(m+2)}{(m-1)!\,(s-m+1)!}$$
$$< \frac{3}{2}\frac{2^{m-1}}{(m-1)!}2s(2s-1)(2s-2)\ldots(s+2).$$

Hence

$$\frac{|\mathrm{Q}_{2m, 2s}(\mu)|}{(2s-1)!} < 3\cdot\frac{2^{m-1}}{(m-1)!}\cdot\frac{s}{(s+1)!} \qquad m \leq s$$

and

$$|\mathrm{Q}_{2s+2, 2s}(\mu)| \leq 3/(4s+6).$$

Consequently

$$\frac{1}{\mathrm{W}a}\cdot\frac{|\mathrm{P}_{2s-1}|}{(2s-1)!} < \frac{3}{s!}\sum_{m=1}^{s+1}\frac{2^{m-1}}{(m-1)!} < \frac{24}{s!}$$

and

$$|\mathrm{P}| \leq 2\omega\sum_{s=1}^{\infty}|\mathrm{P}_{2s-1}|(2\omega t)^{2s-1}/(2s-1)!$$
$$< \sum_{s=1}^{\infty}48\mathrm{W}a\omega(2\omega t)^{2s-1}/s!$$

i.e. the series for P in terms of the sequence P_{2s-1} is absolutely convergent for finite values of t.

Similarly it may be shown that

$$|\text{Gradient of } P_{2s-1}| < (4s+3)|P_{2s-1}|,$$

and that the series for u, v, w, in terms of P_{2s-1}, are absolutely convergent for finite values of t.

The Motion of the Sphere.

§ 6. The velocity of the centre of the sphere is given by

$$V = \sum_{s=0}^{\infty} c_s (2\omega t)^s / s!$$

and from (19)

$$\left.\begin{aligned} c_{2s} &= (-1)^s 3W/\{(2s+1)(2s+3)\} \\ c_{2s+1} &= 0 \end{aligned}\right\}, \qquad s \geq 0.$$

Hence

$$V = \sum_{s=0}^{\infty} (-1)^s \frac{3W}{(2s+1)(2s+3)} \cdot \frac{(2\omega t)^{2s}}{(2s)!}$$

$$= 3W \left\{ \frac{\sin 2\omega t}{(2\omega t)^3} - \frac{\cos 2\omega t}{(2\omega t)^2} \right\}.$$

This may be expressed in a neater form by using the notation

$$S_n(x) = x^{n+1} \left(-\frac{1}{x} \cdot \frac{d}{dx} \right)^n \frac{\sin x}{x}$$

so that

$$V = 3W \, . \, S_1(2\omega t)/(2\omega t)^2.$$

The values of V/W for $2\omega t = 0, 1, 2, \ldots 14$ have been found, and from them the Diagram I constructed, which shows the ratio V/W plotted against that

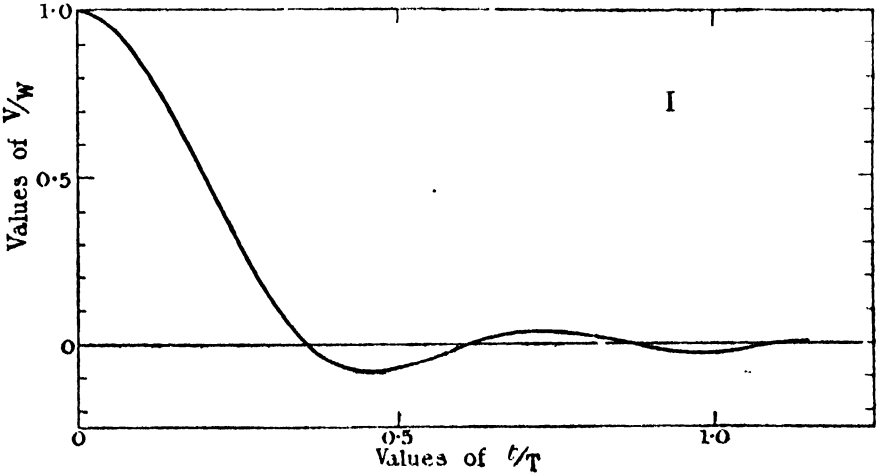


DIAGRAM I.

of t/T, where $T = 2\pi/\omega$ is the time of a revolution of the undisturbed liquid.

It is to be observed that the value of V is oscillatory, but tends rapidly to zero; the sphere oscillates about its final position with continually decreasing amplitude, the period of oscillation tending to a constant value equal to ½T; after one revolution of the liquid the sphere is virtually at rest.

The distance between the initial and final positions of the sphere is given by

$$\int_0^\infty V\,dt = \int_0^\infty 3W\left\{\frac{\sin 2\omega t}{(2\omega t)^3} - \frac{\cos 2\omega t}{(2\omega t)^2}\right\}dt = \frac{1}{2}\cdot\frac{\pi}{2}\cdot\frac{3W}{2\omega} = \frac{3}{16}WT.$$

The quantity WT is the distance the sphere would travel during the time T with its initial velocity of projection, consequently the sphere comes virtually to rest after a time T, at a point 3/16 of this distance from the point of projection.

Velocity of the Liquid along the Axis of the Sphere.

§ 7. Along the axis of the sphere we have $\mu = \pm 1$, and taking first the case $\mu = 1$, we derive from (15)

$$Q_{2m,\,2s} = 0, \qquad m < s,$$

with

$$Q_{2s,\,2s} = (-1)^{s-1}\,3/(4s+2), \qquad Q_{2s+2,\,2s} = (-1)^{s}\,3/(4s+6),$$

so that

$$P_{2s-1} = (-1)^s\frac{3Wa}{2}\left\{\left(\frac{a}{r}\right)^{2s+2}\frac{1}{2s+3} - \left(\frac{a}{r}\right)^{2s}\frac{1}{2s+1}\right\}$$

and on the axis of the sphere

$$\frac{\partial P_{2s-1}}{\partial z} = \frac{\partial P_{2s-1}}{\partial r} = (-1)^{s+1}\frac{3W}{2}\left\{\left(\frac{a}{r}\right)^{2s+3}\frac{2s+2}{2s+3} - \left(\frac{a}{r}\right)^{2s+1}\frac{2s}{2s+1}\right\}.$$

From equations (6) we have

$$w_s = -\partial P_{s-1}/\partial z, \qquad s = 0, 1, 2, \ldots,$$

hence it follows that

$$\left.\begin{aligned} w_{2s+1} &= 0 \\ w_{2s} &= -\frac{\partial P_{2s-1}}{\partial z} = (-1)^s\frac{3W}{2}\left\{\left(\frac{a}{r}\right)^{2s+3}\frac{2s+2}{2s+3} - \left(\frac{a}{r}\right)^{2s+1}\frac{2s}{2s+1}\right\} \end{aligned}\right\} \quad s \geq 0.$$

Since

$$w = \sum_{s=0}^{\infty} w_s(2\omega t)^s/s!$$

we have, on substituting the expressions obtained for the coefficients w_s

$$
\begin{aligned}
w &= \sum_{s=0}^{\infty}(-1)^s \frac{3W}{2}\left\{\left(\frac{a}{r}\right)^{2s+3}\frac{2s+2}{2s+3}-\left(\frac{a}{r}\right)^{2s+1}\frac{2s}{2s+1}\right\}\frac{(2\omega t)^{2s}}{(2s)!} \\
&= \frac{3W}{2}\left(\frac{a}{r}\right)^3 \sum_{s=0}^{\infty}(-1)^s \frac{2s+2}{2s+3}\frac{(2a\omega t/r)^{2s}}{(2s)!} \\
&\qquad\qquad + \frac{3W}{2}\frac{a}{r}\sum_{s=1}^{\infty}(-1)^{s-1}\frac{2s}{2s+1}\frac{(2a\omega t/r)^{2s}}{(2s)!} \\
&= \frac{3W}{2}\left(\frac{a}{r}\right)^3\left\{\cos(2a\omega t/r)-\frac{\sin(2a\omega t/r)}{2a\omega t/r}-2\frac{\cos(2a\omega t/r)}{(2a\omega t/r)^2}\right. \\
&\qquad \left.+2\frac{\sin(2a\omega t/r)}{(2a\omega t/r)^3}\right\}-\frac{3W}{2}\frac{a}{r}\left\{\cos(2a\omega t/r)-\frac{\sin(2a\omega t/r)}{(2a\omega t/r)}\right\} \\
&= \frac{3W}{2}\left\{\frac{a}{r}-\left(\frac{a}{r}\right)^3+2\frac{a}{r}\frac{1}{(2\omega t)^2}\right\}S_1\left(\frac{2a\omega t}{r}\right), \qquad (21)
\end{aligned}
$$

which, on $r = a$, agrees with the expression obtained for the velocity of the sphere.

The case $\mu = -1$ gives the same expression for w, and, on the axis of the sphere,

$$u = v = 0.$$

Perhaps the clearest way to discuss the function w is to examine the variation with r for constant values of t. A table has been constructed and from it the Diagram II formed, in which the ratio w/W is plotted against that of r/a; the curves are numbered 0, 1, 2, ... 10, and represent w/W at

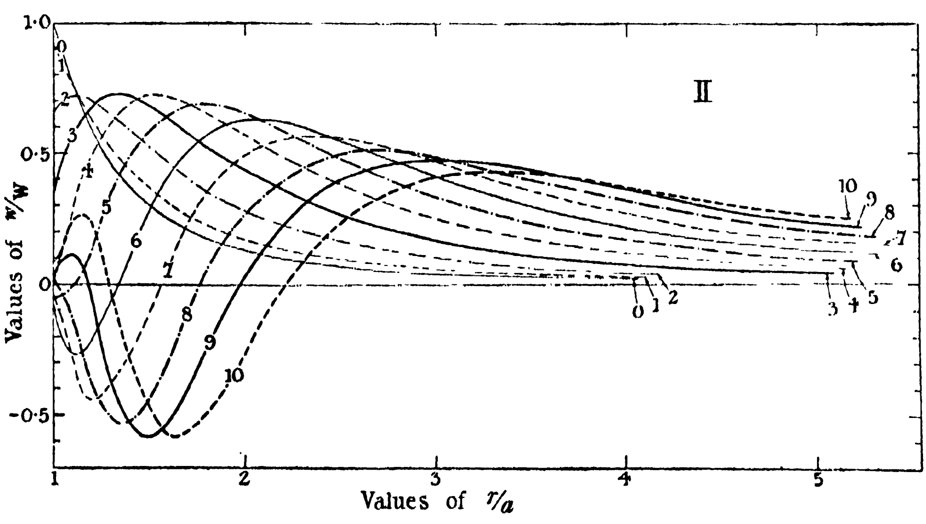


Diagram II.

points on the axis of the sphere for $2\omega t = 0, 1, 2, \ldots 10$ respectively. It may be noted that $2\omega t = 10$ corresponds roughly to $t = 0{\cdot}8\mathrm{T}$.

A further Diagram III has been drawn showing the velocities at times corresponding to T, 2T, 3T, respectively.

It may be noted that the function $\mathrm{S}_1(2a\omega t/r)$ oscillates between bounds in the neighbourhood of ± 1. Referring now to the Diagrams II and III and to the expression for w given in (21), it is seen that w is always less than W. As r increases, w is ultimately positive and tends to zero. At intermediate points the velocity may be positive or negative, the disturbance being a maximum in the neighbourhood of $r = 1{\cdot}7a$. Here it is to be noted, however, that as t increases the velocity gradient along the axis becomes large, due to the fact that in $\partial w/\partial r$ there is a factor $2\omega t$. This point is clearly illustrated in Diagram III, which indicates the state of the liquid as t becomes large; the changes of velocity are observed to be contained within upper and lower bounds, which are $\pm\frac{3}{2}\mathrm{W}\{a/r-(a/r)^3\}$, since for t large we have approximately

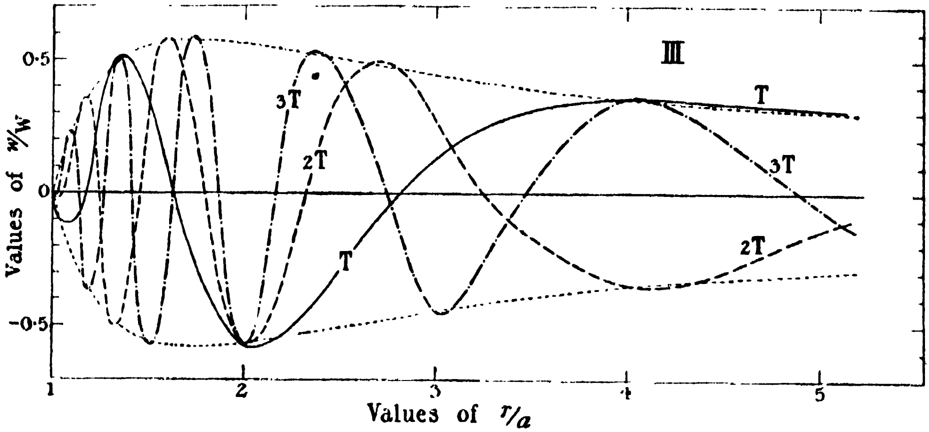


DIAGRAM III.

$$w = \frac{3}{2}\mathrm{W}\left(\frac{a}{r}-\frac{a^3}{r^3}\right)\mathrm{S}_1\left(\frac{2a\omega t}{r}\right).$$

Velocity of the Liquid on the Equatorial Plane of the Sphere.

§ 8. The equatorial plane of the sphere is the plane defined by $\mu = 0$. Here we have

$$u = v = 0,$$

so that the velocity is wholly parallel to the axis of rotation. We have

$$\left(\frac{\partial \mathrm{P}'_{2s-1}}{\partial z}\right)_{\mu=0} = \frac{1}{r}\left(\frac{\partial \mathrm{P}'_{2s-1}}{\partial \mu}\right)_{\mu=0}$$

and from (14)

$$\left(\frac{\partial Q_{2m,2s}}{\partial\mu}\right)_{\mu=0} = (-1)^{s-m}\frac{3}{2}\cdot\frac{1.3.5\ldots.(2s-1)}{2^s(m-1)!\,(s-m+1)!}\left\{\frac{2}{2m+1}-1\right\} \qquad s,\ m>0$$

$$= (-1)^{s-m+1}\frac{3}{2}\cdot\frac{2m-1}{2m+1}\cdot\frac{1.3.5\ldots.(2s-1)}{2^s(m-1)!\,(s-m+1)!}$$

so that

$$\left(\frac{\partial P_{2s-1}}{\partial z}\right)_{\mu=0} = (-1)^s\frac{3W}{2}\cdot\frac{1.3.5\ldots.(2s-1)}{2^s}\times$$

$$\left\{\frac{1}{3}\cdot\frac{1}{s!}\left(\frac{a}{r}\right)^3-\frac{3}{5}\frac{1}{(s-1)!\,1!}\left(\frac{a}{r}\right)^5+\frac{5}{7}\frac{1}{(s-2)!\,2!}\left(\frac{a}{r}\right)^7-\ldots\right.$$

$$\ldots-(-1)^m\frac{2m-1}{2m+1}\frac{1}{(s-m+1)!\,(m-1)!}\left(\frac{a}{r}\right)^{2m+1}+\ldots$$

$$\left.\ldots+(-1)^s\frac{2s+1}{2s+3}\cdot\frac{1}{s!}\left(\frac{a}{r}\right)^{2s+3}\right\}$$

$$= (-1)^s\frac{3W}{2}\cdot\frac{1.3.5\ldots.(2s-1)}{2^s\,.\,s!}\times$$

$$\left\{\frac{1}{3}\left(\frac{a}{r}\right)^3-{}^sC_1\frac{3}{5}\left(\frac{a}{r}\right)^5+{}^sC_2\frac{5}{7}\left(\frac{a}{r}\right)^7-\ldots\right.$$

$$\left.-(-1)^{m}\,{}^sC_{m-1}\frac{2m-1}{2m+1}\left(\frac{a}{r}\right)^{2m+1}+\ldots+(-1)^s\frac{2s+1}{2s+3}\left(\frac{a}{r}\right)^{2s+3}\right\}. \tag{22}$$

The quantity within the brackets, on putting $a/r = \lambda$, becomes

$$\frac{1}{3}\lambda^3-{}^sC_1\frac{3}{5}\lambda^5+{}^sC_2\frac{5}{7}\lambda^7-\ldots+(-1)^s\frac{2s+1}{2s+3}\lambda^{2s+3}$$

$$=\int_0^\lambda\lambda^2(1-\lambda^2)^s\,d\lambda-2s\int_0^\lambda\lambda^4(1-\lambda^2)^{s-1}\,d\lambda$$

$$=\lambda^3(1-\lambda^2)^s-2\int_0^\lambda\lambda^2(1-\lambda^2)^s\,d\lambda. \tag{23}$$

Hence

$$\left(\frac{\partial P_{2s-1}}{\partial z}\right)_{\mu=0} = (-1)^s\frac{3W}{2}\cdot\frac{1.3.5\ldots.(2s-1)}{2^s\,.\,s!}\times$$

$$\left\{\lambda^3(1-\lambda^2)^s-2\int_0^\lambda\lambda^2(1-\lambda^2)^s\,d\lambda\right\},$$

so that, on the equatorial plane,

$$w=\sum_{s=0}^{\infty}w_s(2\omega t)^s/s!$$

$$=-\tfrac{1}{2}W\lambda^3-\sum_{s=1}^{\infty}(-1)^s\frac{3W}{2}\cdot\frac{1.3.5\ldots.(2s-1)}{2^s\,.\,s!}\times$$

$$\left\{\lambda^3(1-\lambda^2)^s-2\int_0^\lambda\lambda^2(1-\lambda^2)^s\,d\lambda\right\}\frac{(2\omega t)^{2s}}{(2s)!}.$$

Further,

$$\sum_{s=1}^{\infty}(-1)^s\frac{1\,.\,3\,.\,5\,\ldots\,(2s-1)}{2^s\,.\,s\,!}(1-\lambda^2)^s\frac{(2\omega t)^{2s}}{(2s)\,!}$$

$$=\sum_{s=1}^{\infty}(-1)^s\frac{\{2\omega t\sqrt{(1-\lambda^2)}\}^{2s}}{2^{2s}(s\,!)^2}=\mathrm{J}_0\{2\omega t\sqrt{(1-\lambda^2)}\}-1,$$

$\mathrm{J}_0\{2\omega t\sqrt{(1-\lambda^2)}\}$ being the Bessel function of zero order. Hence we have

$$w=-\tfrac{1}{2}\mathrm{W}\lambda^3-\tfrac{3}{2}\mathrm{W}\Big[\lambda^3\mathrm{J}_0\{2\omega t\sqrt{(1-\lambda^2)}\}$$

$$-\lambda^3-2\int_0^\lambda[\lambda^2\mathrm{J}_0\{2\omega t\sqrt{(1-\lambda^2)}\}-\lambda^2]\,d\lambda\Big]$$

$$=-\tfrac{3}{2}\mathrm{W}\Big[\lambda^3\mathrm{J}_0\{2\omega t\sqrt{(1-\lambda^2)}\}-2\int_0^\lambda\lambda^2\mathrm{J}_0\{2\omega t\sqrt{(1-\lambda^2)}\}\,d\lambda\Big].\qquad(24)$$

This is the most compact expression which has been obtained for the velocity at points on the equatorial plane; it is convenient in that it enables us to obtain an approximation to the value of w.

When λ is small, the value of the integral is approximately $\frac{1}{3}\lambda^3\mathrm{J}_0(2\omega t)$, provided $2\omega t$ is not large, which gives w in the form

$$-\tfrac{1}{2}\mathrm{W}\lambda^3\mathrm{J}_0(2\omega t).$$

To approximate to the value of w for t large, we use the well-known fact that, for x large, $\mathrm{J}_0(x)$ is approximately

$$\sqrt{\left(\frac{2}{\pi x}\right)}\cos\left(x-\frac{\pi}{4}\right)$$

so that, provided $\sqrt{(1-\lambda^2)}$ is finite and $2\omega t$ sufficiently large, w becomes

$$-\frac{3\mathrm{W}\lambda^3}{2}\sqrt{\left\{\frac{2}{\pi\,2\omega t\sqrt{(1-\lambda^2)}}\right\}}\cos\left\{2\omega t\sqrt{(1-\lambda^2)}-\frac{\pi}{4}\right\}$$

$$+3\mathrm{W}\int_0^\lambda\lambda^2\sqrt{\left\{\frac{2}{\pi\,2\omega t\sqrt{(1-\lambda^2)}}\right\}}\cos\left\{2\omega t\sqrt{(1-\lambda^2)}-\frac{\pi}{4}\right\}d\lambda,$$

and, on examination, we see that ultimately the integral is of a higher order of small quantities, so that w is approximately

$$-\frac{3\mathrm{W}\lambda^3}{2}\sqrt{\left\{\frac{2}{\pi\,2\omega t\sqrt{(1-\lambda^2)}}\right\}}\cos\left\{2\omega t\sqrt{(1-\lambda^2)}-\frac{\pi}{4}\right\}.$$

It follows, therefore, that there is ultimately an oscillation about the equatorial plane, the amplitude at any point continually decreasing and tending to zero, the period, however, tending to the constant value

$$\frac{2\pi}{2\omega\sqrt{(1-\lambda^2)}}=\frac{1}{2}\cdot\frac{\mathrm{T}}{\sqrt{\{1-(a/r)^2\}}}.$$

Again, since $\partial w/\partial r$ contains a factor $(2\omega t)^{\frac{1}{2}}$, the velocity gradient ultimately increases without limit.

To examine further the equatorial velocity, we put the expression (24) into a power series in λ. To obtain this form we return to the expression for $(\partial P_{2s-1}/\partial z)_{\mu=0}$ given by (22) and consider the coefficient of λ^{2m+3} for $m > 0$ in

$$\sum_{s=1}^{\infty}\left(\frac{\partial P_{2s-1}}{\partial z}\right)_{\mu=0}\frac{(2\omega t)^{2s}}{2s\,!}.$$

It is

$$(-1)^m\frac{2m+1}{2m+3}\cdot\frac{3W}{2}\sum_{s=m}^{\infty}(-1)^{s}\,{}^{s}C_m\frac{(2\omega t)^{2s}}{2^{2s}(s\,!)^2}$$

$$=\frac{2m+1}{2m+3}\cdot\frac{1}{2^m m\,!}\cdot\frac{3W}{2}(2\omega t)^m\sum_{p=0}^{\infty}(-1)^p\frac{(2\omega t)^{2p+m}}{2^{2p+m}p\,!\,(p+m)\,!}$$

$$=\frac{3W}{2}\cdot\frac{2m+1}{2m+3}\cdot\frac{1}{2^m\,m\,!}(2\omega t)^m J_m(2\omega t),$$

and the coefficient of λ^3 is

$$-\tfrac{1}{2}W\,\{1-J_0(2\omega t)\}.$$

Hence w is given by

$$-\frac{3}{2}W\left\{\frac{1}{3}\lambda^3J_0(2\omega t)+\frac{3}{5}\cdot\frac{1}{2}\lambda^5\,2\omega tJ_1(2\omega t)+\frac{5}{7}\cdot\frac{1}{2\,.\,4}\lambda^7(2\omega t)^2J_2(2\omega t)+\dots\right.$$
$$\left.\dots+\frac{2m+1}{2m+3}\,\frac{1}{2\,.\,4\,.\,6\,\dots\,2m}\lambda^{2m+3}(2\omega t)^mJ_m(2\omega t)+\dots\right\}.$$

From this expression we are able to construct a Table giving the values of w for $2\omega t = 0, 1, 2, \dots 10$, and to graph the function w for constant values of t (Diagram IV). The Diagrams II, III, IV are all drawn to the

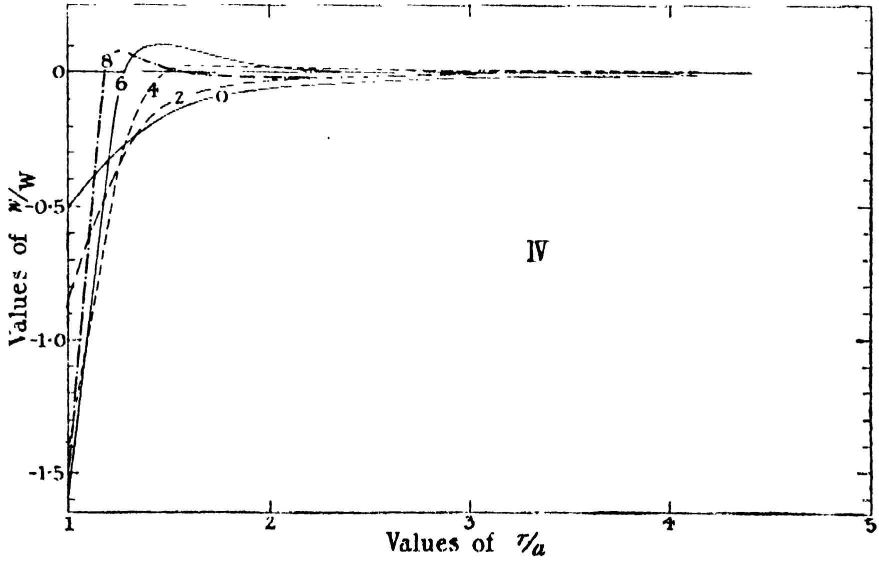


DIAGRAM IV.

same scale for purposes of comparison, but, in the last case, only the curves corresponding to $2\omega t = 0, 2, 4, 6, 8$ are shown, since the variation is so small.

An examination of Diagram IV confirms the results already given; it also indicates that the disturbance on the equatorial plane is practically confined to the liquid in the immediate neighbourhood of the sphere.

On the sphere itself, putting $\lambda = 1$ in (23),

$$\left(\frac{\partial P_{2s-1}}{\partial z}\right)_{\mu=0} = (-1)^{s+1}\frac{3W}{2}\frac{1.3.5.\ldots(2s-1)}{2^s.s!}2\cdot\frac{1}{2s+3}\cdot\frac{2s}{2s+1}\cdot\frac{2s-2}{2s-1}\cdots\frac{2}{3}$$

$$= (-1)^{s+1}3W/\{(2s+1)(2s+3)\},$$

so that, on the equator of the sphere,

$$w = -\frac{W}{2} - 3W\sum_{s=1}^{\infty}(-1)^{s+1}\frac{1}{(2s+1)(2s+3)}\cdot\frac{(2\omega t)^{2s}}{(2s)!}$$

$$= -\tfrac{3}{2}W + 3W \,.\, S_1(2\omega t)/(2\omega t)^2,$$

which rapidly tends to the limiting value $-\frac{3}{2}W$.

Velocity of the Liquid at Points on the Sphere.

§ 9. To discuss the velocity of the liquid at points on the surface of the sphere, we transform the dynamical equations (3) into spherical polar co-ordinates, and obtain

$$\frac{\partial u'}{\partial t} - 2\omega v'\cos\theta = -\frac{\partial P}{r\partial\theta},$$

$$\frac{\partial v'}{\partial t} + 2\omega w'\sin\theta + 2\omega u'\cos\theta = 0, \tag{25}$$

$$\frac{\partial w'}{\partial t} - 2\omega v'\sin\theta = -\frac{\partial P}{\partial r}, \tag{26}$$

where w' coincides with w for $\theta = 0$ and $-u'$ coincides with w on the equator.

On the sphere, we have

$$(w')_{r=a} = V\cos\theta = 3W\cos\theta\left\{\frac{\sin 2\omega t}{(2\omega t)^3} - \frac{\cos 2\omega t}{(2\omega t)^2}\right\}$$

and from (20)

$$\left(\frac{\partial P_{2s-1}}{\partial r}\right)_{r=a} = (-1)^s W\left\{\frac{3}{2}\cos^{2s-1}\theta\sin^2\theta - \frac{3\cos\theta}{(2s+1)(2s+3)}\right\} \quad s > 0,$$

so that

$$\left(\frac{\partial \mathrm{P}}{\partial r}\right)_{r=a} = 2\omega \sum_{s=0}^{\infty} \left(\frac{\partial \mathrm{P}s}{\partial r}\right)_{r=a} \frac{(2\omega t)^s}{s\,!}$$

$$= 2\omega \sum_{s=1}^{\infty} (-1)^s\, \mathrm{W} \left\{\frac{3}{2}\cos^{2s-1}\theta \sin^2\theta - \frac{3\cos\theta}{(2s+1)(2s+3)}\right\} \frac{(2\omega t)^{2s-1}}{(2s-1)!}$$

$$= -2\omega \mathrm{W}\left[\frac{3}{2}\sin^2\theta \sin(2\omega t\cos\theta)\right.$$

$$\left. + 3\cos\theta\left\{\frac{\sin 2\omega t}{(2\omega t)^2} + \frac{3\cos 2\omega t}{(2\omega t)^3} - \frac{3\sin 2\omega t}{(2\omega t)^4}\right\}\right].$$

Hence from (26)

$$(v')_{r=a} = -\tfrac{3}{2}\,\mathrm{W}\sin\theta\sin(2\omega t\cos\theta),$$

and from (25)

$$(-u')_{r=a} = -\tfrac{3}{2}\,\mathrm{W}\sin\theta\cos(2\omega t\cos\theta) + 3\,\mathrm{W}\sin\theta\ .\ \mathrm{S}_1(2\omega t)/(2\omega t)^2. \qquad (27)$$

The expression for $(w')_{r=a}$ has already been treated; it is simply the radial component of the velocity of the sphere.

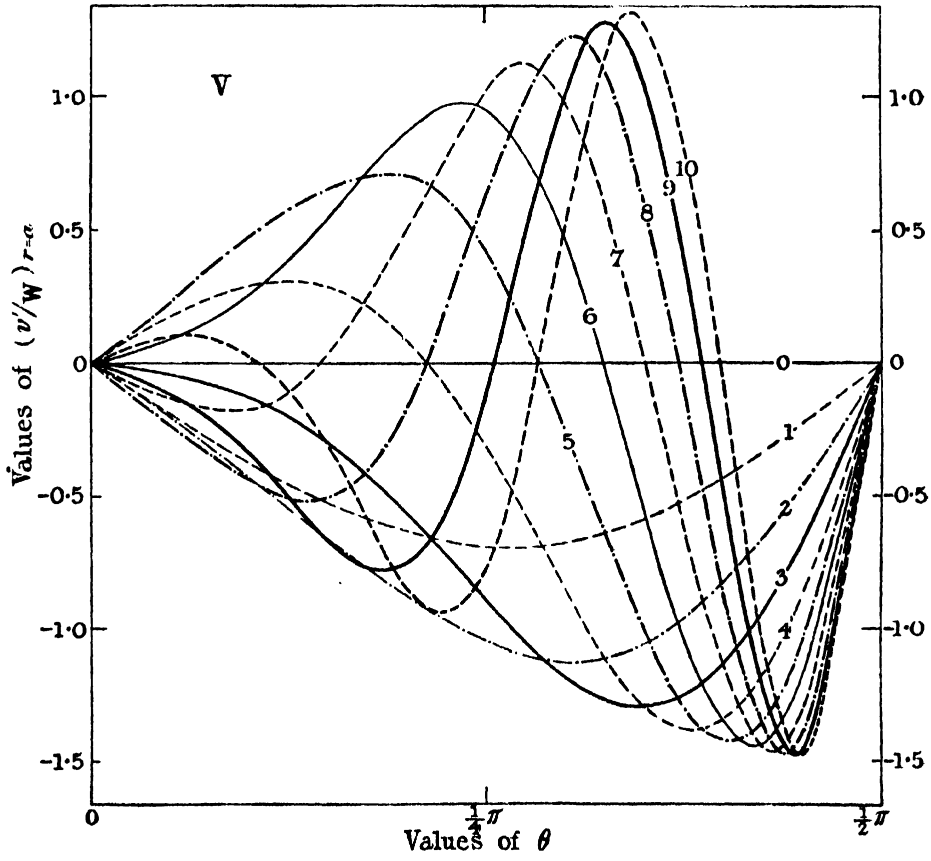


DIAGRAM V.

Considering the expression for $(v')_{r=a}$, a Table may be formed giving its value for θ between 0 and $\frac{1}{2}\pi$ at times corresponding to $2\omega t = 0, 1, 2, \ldots 10$ respectively, and from it the Diagram V constructed. The values for θ from $\frac{1}{2}\pi$ to π are those from 0 to $\frac{1}{2}\pi$, repeated in the reverse order with change of sign.

It is observed that the transverse velocity, $(v')_{r=a}$, oscillates between bounds $\pm\frac{3}{2}\mathrm{W}\sin\theta$, which are independent of the time.

The velocity along a meridian, coinciding with w on the equator, is given by (27). The variation of this quantity is indicated by the Diagram VI, which is constructed similarly to V. The values for θ from $\frac{1}{2}\pi$ to π are those from 0 to $\frac{1}{2}\pi$ in the reverse order.

The value of $(-u')_{r=a}$ oscillates between bounds in the neighbourhood of $\pm\frac{3}{2}\mathrm{W}\sin\theta$, approaching these bounds more nearly as t increases; the

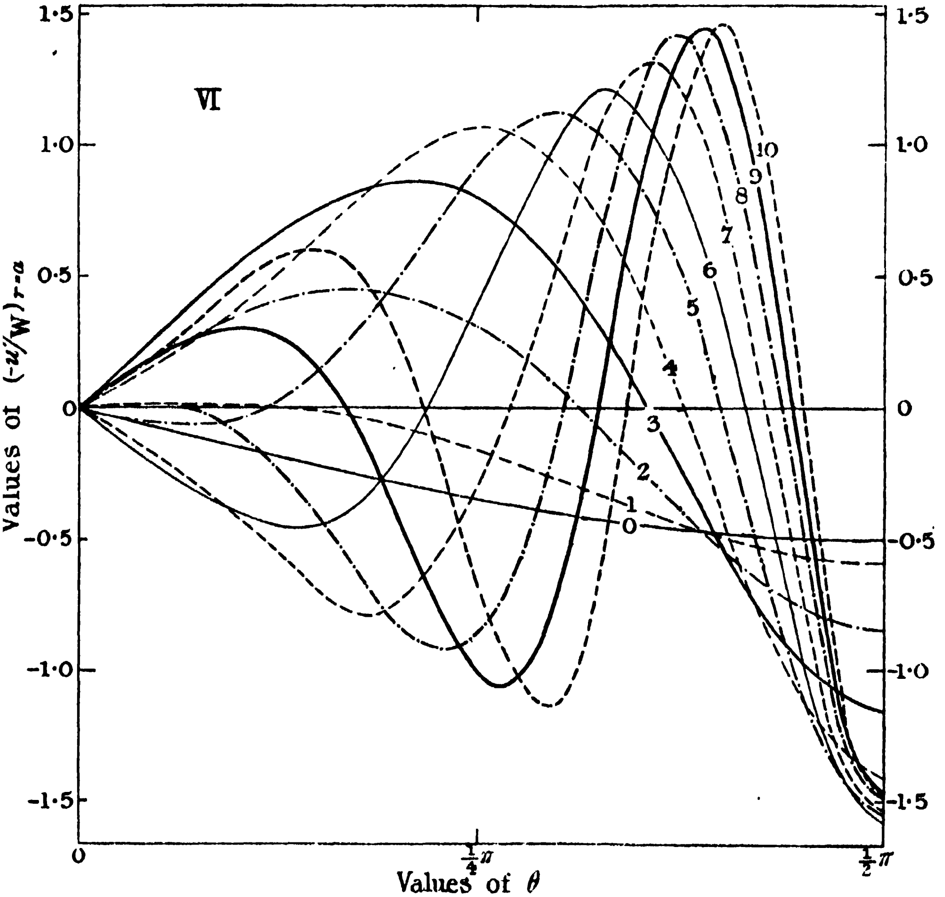


DIAGRAM VI.

irregularity is due to the term $3\mathrm{W}\sin\theta\,.\,\mathrm{S}_1\,(2\omega t)/(2\omega t)^2$, which rapidly tends to zero.

Hence, for t large, the velocity of the liquid tangential to the sphere at a point of co-latitude θ, has components

$$-\tfrac{3}{2}\,\mathrm{W}\sin\theta\cos(2\omega t\cos\theta), \qquad -\tfrac{3}{2}\,\mathrm{W}\sin\theta\sin(2\omega t\cos\theta),$$

which combine into $\tfrac{3}{2}\mathrm{W}\sin\theta$ in a direction making an angle α, measured in the negative direction from the meridian (θ increasing) through the point where $\alpha = 2\omega t\cos\theta$.

Hence, ultimately, the velocity of the liquid tangential to the sphere at any point is constant in magnitude and equal to $\tfrac{3}{2}\mathrm{W}\sin\theta$, but changes its direction with constant angular speed $2\omega\cos\theta$, which is a maximum at the poles and zero at the equator.

A point to notice, however, is the fact that $(\partial v'/\partial\theta)_{r=a}$ contains the factor $2\omega t$, and therefore ultimately increases without limit. This means that we reach a stage beyond which the components of vorticity can no longer be regarded as small. A similar deduction is obtained from $\partial w/\partial r$ over the equatorial plane, which contains the factor $(2\omega t)^{\frac{1}{2}}$. A physical restriction to the solution of the problem is thus introduced.

On the Mutual Threading of Vortex Rings.

By W. M. Hicks, F.R.S.

(Received June 26, 1922.)

Although the fact that two circular vortices are capable of threading one another in permanent succession has long been known, I do not know that any attempt has been made at a discussion of the general conditions. The following pages contain a contribution to such a discussion for the case of thin circular filaments. The nature of a ring is given when its circulation and volume are known. Its configuration at any time is given when its aperture *or* the radius of the cross-section of the filament is known. In the case of two rings, the nature of their combination is defined if their two apertures at the instant when they are co-planar are each known. This configuration will be called their standard position. With two such rings, the mean area of the two apertures, supposed weighted with their circulations, remains constant throughout the motion. They can therefore also be defined by the radius of this mean area and the ratio of the two apertures in the standard position. This latter method is the more convenient. It is found possible to determine the relative paths of one as seen from the other, and to obtain the conditions of permanent union for any two given rings, but the complete solution in terms of the time, and the actual paths in the fluid, are not expressible in general terms. Such can only be arrived at by numerical quadrature for cases in which the various coefficients involved have their numerical values given. Any special case can thus be solved, but naturally the process is very laborious.

In the corresponding problem in two dimensions, however, a complete solution is attainable. For this reason, and because it illustrates the general method, the problem of two like and equal two-dimensional pairs (rectilinear filaments) is first touched upon. The more complete discussion for all the possible combinations is extremely interesting, especially when the pairs have opposite circulations, but is here omitted, in view of the greater physical importance of the theory of the ring.

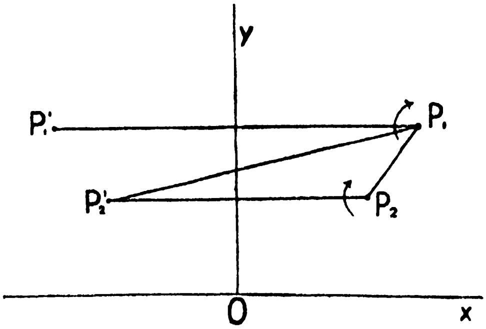


FIG. 1.

I.—*Parallel Straight Vortices.*

In the figure, P_1, P_1', P_2, P_2' denote the positions of the four filaments. Let x_1, y_1, x_2, y_2 denote the co-ordinates of P_1, P_2. Further, let x, y denote the co-ordinate of P_1 referred to P_2, *i.e.*,

$$x = x_1 - x_2, \qquad y = y_1 - y_2,$$

The stream functions of the pairs at a point ξ, η are given by $\mu_1\chi_1/4\pi$, $\mu_2\chi_2/4\pi$, where

$$\chi_1 = -\log\left(\frac{PP_1}{PP_1'}\right)^2 = -\log\frac{(\xi-x_1)^2+(\eta-y_1)^2}{(\xi+x_1)^2+(\eta-y_2)^2}. \tag{1}$$

It should be noted that the values of χ_1 at P_2 and of χ_2 at P_1 are the same.

Write $(PP_1/PP_1')^2 = 1-f_1^2$ where

$$f_1^2 = \frac{4\xi x_1}{(\xi+x_1)^2+(\eta-y_1)^2}$$

and let f denote the common value of f_1 at P_2 or of f_2 at P_1. Then, putting $\mu_1 = \mu_2 = \mu$

$$\left.\begin{aligned} \frac{dx_1}{dt} &= -\frac{\mu}{4\pi}\left.\frac{\overline{d\chi_2}}{dy}\right|_1 = \frac{\mu}{\pi}\frac{f^3}{4x_1x_2}y\frac{d\chi}{df} \\ \frac{dx_2}{dt} &= \ldots \qquad\qquad = -\frac{\mu}{\pi}\frac{f^3}{4x_1x_2}y\frac{d\chi}{df} \end{aligned}\right\}, \tag{2}$$

whence

$$x_1+x_2 = \text{const.} = A,$$

$$\frac{dx}{dt} = \frac{\mu}{4\pi x_1x_2}\frac{f^4}{1-f^2}y, \tag{3}$$

and

$$f^2 = \frac{A^2-x^2}{A^2+y^2}.$$

Also

$$\left.\begin{aligned} \frac{dy_1}{dt} &= \frac{\mu}{4\pi x_1}+\frac{\mu}{4\pi}\left.\frac{\overline{d\chi_2}}{dx}\right|_1 = \frac{\mu}{4\pi x_1}+\frac{\mu}{4\pi}\left(\frac{1}{2x_1}-\frac{x_1+x_2}{4x_1x_2}f^2\right)f\frac{d\chi}{df} \\ \frac{dy_2}{dt} &= \ldots \qquad\qquad = \frac{\mu}{4\pi x_2}+\frac{\mu}{4\pi}\left(\frac{1}{2x_2}-\frac{x_1+x_2}{4x_1x_2}f^2\right)f\frac{d\chi}{df} \end{aligned}\right\}. \tag{4}$$

Hence

$$\frac{dy}{dt} = -\frac{\mu}{4\pi x_1x_2}\frac{1}{1-f^2}x, \tag{5}$$

$$\therefore \qquad \frac{dx}{dy} = -f^4\frac{y}{x} = -\frac{y}{x}\left(\frac{A^2-x^2}{A^2+y^2}\right)^2,$$

The integral of this is

$$\frac{1}{A^2+y^2}-\frac{1}{A^2-x^2} = B. \tag{6}$$

2. We shall define a given configuration by the apertures when the two pairs are co-planar, say $2a$, $2b$, $a > b$. Then

$$A = a+b, \qquad B = \frac{1}{(a+b)^2}-\frac{1}{4ab}.$$

It follows from (6) that

$$y^2 = \frac{(a+b)^4}{(a-b)^2}\frac{(a-b)^2-x^2}{x^2+(a+b)^2\{4ab/(a-b)^2-1\}}.$$

If, then, $4ab > (a-b)^2$, y is always finite, or the filaments revolve round each other in permanent union. If $4ab < (a-b)^2$, one pair simply passes through the aperture of the other, and then separates to infinity. The above condition gives for permanent union

$$a > (3+2\sqrt{2})b, \qquad > 5{\cdot}8284b,$$
$$b < 0{\cdot}17157a.$$

Equation (6) may also be written

$$x^2 = (a+b)^2\frac{(a^2-b^2)^2+\{(a-b)^2-4ab\}y^2}{(a+b)^4+(a-b)^2y^2}.$$

In this put

$$\frac{(a^2-b^2)^2}{4ab-(a-b)^2} = \alpha^2, \qquad \frac{(a+b)^4}{(a-b)^2} = \beta^2.$$

Then

$$x^2 = \frac{(a+b)^4}{\alpha^2}\,\frac{\alpha^2-y^2}{\beta^2+y^2}. \tag{7}$$

3. *Period of Revolution.*—From (5) since $4r_1r_2 = A^2 - x^2$,

$$\frac{dy}{dt} = -\frac{\mu}{\pi}\,\frac{A^2+y^2}{x^2+y^2}\cdot\frac{x}{A^2-x^2}.$$

By (7) it can be shown that

$$\frac{dy}{dt} = -\frac{\mu}{\pi}\,\frac{(a-b)^2}{4ab\alpha}\left(\frac{\beta^2+y^2}{A^2+y^2}\right)^2\sqrt{\frac{\alpha^2-y^2}{\beta^2+y^2}}.$$

Put $y = \alpha\cos\phi$ and $k^2 = \alpha^2/(\alpha^2+\beta^2) = (a-b)^4/16a^2b^2$ so that

$$k = \frac{(a-b)^2}{4ab}.$$

Then

$$\frac{dt}{d\phi} = \frac{\pi(a+b)^2}{\mu k(1-k)}\left\{\sqrt{(1-k^2\sin^2\phi)} - \frac{2(1-k)}{\sqrt{(1-k^2\sin^2\phi)}} + \frac{(1-k)^2}{(1-k^2\sin^2\phi)^{3/2}}\right\}$$

whence integrating

$$t = \frac{2\pi(a+b)^2}{\mu k}\left\{\frac{1}{k'^2}\,E(\phi) - F(\phi) - \frac{1}{2}\,\frac{k^2}{1+k}\,\frac{\sin\phi\cos\phi}{\sqrt{(1-k^2\sin^2\phi)}}\right\}. \tag{8}$$

The period of a complete revolution is therefore

$$T = \frac{8\pi(a+b)^2}{\mu k}\left(\frac{1}{k'^2}\,E - F\right) = \frac{8\pi(a+b)^2}{\mu}\,\frac{dF}{dk}. \tag{8a}$$

Thus, for a given ratio of apertures, the period varies as the square of their linear values.

Expressed in terms of the angle ϕ, the relative co-ordinates are

$$\left.\begin{aligned} x &= (a+b)\sqrt{\,.\,k(1-k)}\;\frac{\sin\phi}{\sqrt{(1-k^2\sin^2\phi)}} \\ y &= (a+b)\sqrt{\left(\frac{k}{1-k}\right)}\cos\phi \end{aligned}\right\}. \tag{9}$$

4. *Motion in the Fluid.*—The relative motion is completely determined by the preceding results. To determine the actual motion in the fluid itself, it is only necessary to determine further the motion of the centre of gravity in the direction of translation.

If Y denote its position at time t,

$$2\mathrm{Y} = y_1 + y_2,$$

$$2\frac{d\mathrm{Y}}{dt} = \frac{\mu(x_1 + x_2)}{4\pi x_1 x_2}(1+f^2) \text{ from (4, 5)}$$

and

$$2\frac{d\mathrm{Y}}{dy} = -\mathrm{A}(1-f^4)\frac{1}{x}.$$

Substituting for x, y their values in terms of ϕ from (9)

$$2\frac{d\mathrm{Y}}{d\phi} = \frac{a+b}{1-k}\left\{-\sqrt{(1-k^2\sin^2\phi)} + \frac{(1-k)^2}{(1-k^2\sin^2\phi)^{3/2}}\right\},$$

$$2\mathrm{Y} = (a+b)\frac{k}{k'^2}\left\{2\mathrm{E}(\phi) - k(1-k)\frac{\sin\phi\cos\phi}{\sqrt{(1-k^2\sin^2\phi)}}\right\}.$$

Hence the total distance traversed during a complete revolution is

$$4(a+b)\frac{k}{k'^2}\mathrm{E}$$

and is independent of the strength of the vortices.

The path of a filament is then given by

$$x' = x_1 = \tfrac{1}{2}(a+b+x),$$

$$y' = \mathrm{Y} + \tfrac{1}{2}y.$$

The mean velocity of propagation of the system through the fluid is

$$\frac{\mathrm{Y}}{\mathrm{T}} = \frac{\mu}{2\pi(a+b)}\frac{k^2\mathrm{E}}{\mathrm{E}-k'^2\mathrm{F}}.$$

Since $k'^2\mathrm{F} > (1-k^2)\mathrm{E}$, this mean velocity is greater than $\mu/[2\pi(a+b)]$, the velocity of progression of a single pair by themselves of aperture $a+b$. Also, as $k^2\mathrm{E}/(\mathrm{E}-k'^2\mathrm{F})$ is less than 2, this mean velocity is less than twice the latter, as, indeed, is evident from general considerations. In fact, the mathematically limiting case corresponds to that in which, while $a-b$ is finite, $a+b$ is infinite, when the two velocities become equal.

5. *Graphical Construction.*—The co-ordinates of the relative orbit may be written

$$y = c\cos\phi,$$

$$x = \frac{1-k}{k}c\frac{c\sin\phi}{\sqrt{(c^2/k^2 - c^2\sin^2\phi)}},$$

where $c = (a+b)\sqrt{(k/.1-k)}$. They suggest the following graphical method for tracing the curves.

Take two rectangular axes Ox, Oy. Draw three concentric circles (centre O) of radii, c, $c(1-k)/k$, and c/k. The last is clearly the largest and the second is larger than the first so long as $k < \frac{1}{2}$. Take any point P on the first and draw PM perpendicular to Ox cutting the outer circle in R. Join OR cutting the second circle in Q, and draw QT perpendicular to OQ, meeting Oy in T. Then

$$\text{angle POT} = \phi$$

$$\text{PM} = y$$

$$\text{QT} = x$$

The construction is simplified if $k = \frac{1}{2}$ since then the first and second circles coincide, and the third has twice their radius.

6. The foregoing results are illustrated by application to two special cases of $k = \sin 10^\circ$ and $\sin 60^\circ$. These correspond to ratios of aperture in the standard co-planar positions of $b/a = 0{\cdot}444$ and $0{\cdot}189$. The relative paths are given in fig. 2 when the sum $a+b$ is the same. Curve A gives that for $b/a = 0{\cdot}189$ (that nearer the limiting case) and curve B that for $b/a = 0{\cdot}444$. The dots on the curves give the positions at equal intervals of time, equal to one-sixteenth of the total period. They serve to show how the velocity varies along the path. It is especially noticeable how in the case of small b/a—curve A—the motion becomes very slow when the apertures approach equality. The corresponding paths in the fluid are given by figs. 3 and 4 for the right hand components, the thick line curve belonging to one component and the thin to the other. The paths described by the left hand components correspond (right and left reversed) at distances on the left of those in the figures respectively equal to 1·47 and 2·60 times the distance 4 ... 4 in the co-planar position. The dots on these curves again give positions at equal intervals. The curves are scrolls, each curve intersecting itself. It might be thought that possibly, for some critical state, the small loop might degenerate to a cusp, followed by states with sinuous paths. The condition would be given by making $dy_1/dt = 0$ when $y = 0$. It would be found that this corresponds to the critical state for permanent union. In other words the paths are always scrolls when in permanent union and sinuous when not.

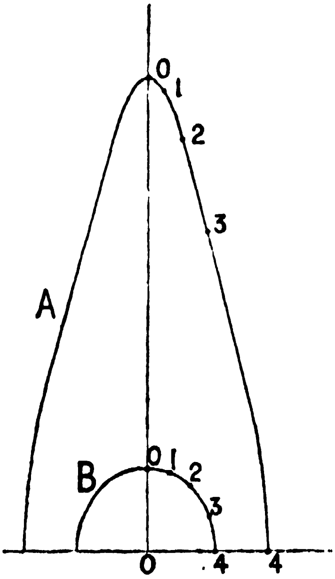


Fig. 2.

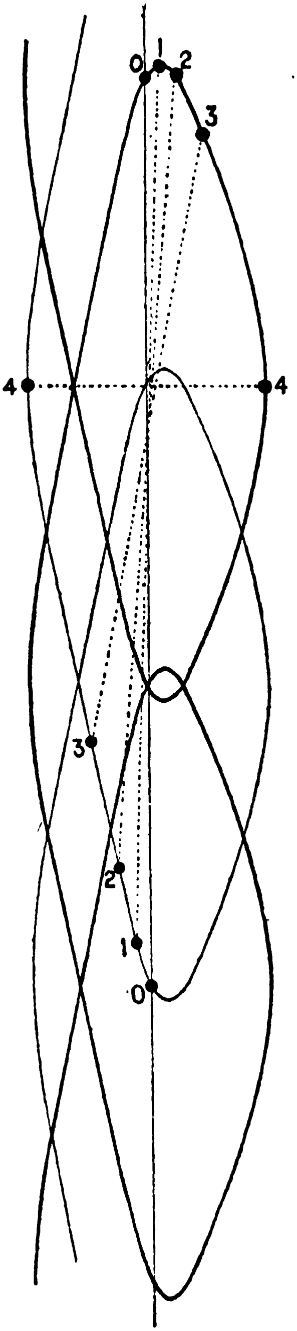


Fig. 3.

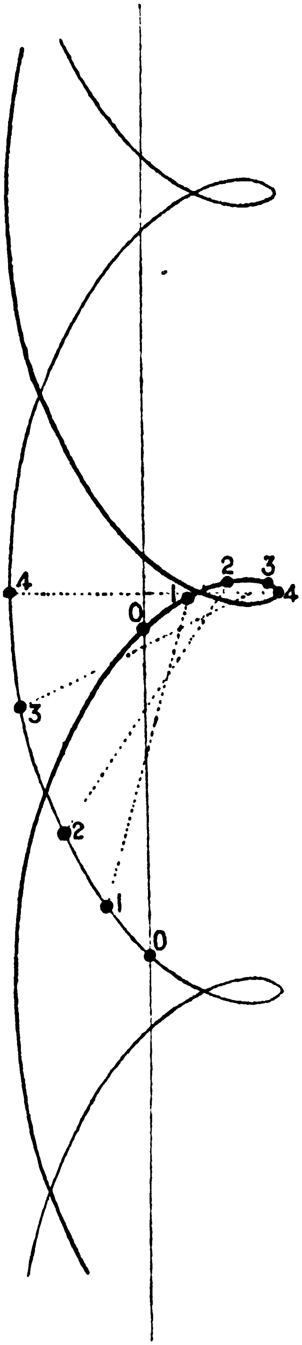


Fig. 4.

II. *Two Circular Rings.*

7. We now pass to our proper problem of the motion of two circular vortex filaments with the same straight axis. In the general case the circulations will be different, say μ_1, μ_2. Let x_1, x_2 denote the radii of their circular axes at any time and y_1, y_2 the distances of their planes from a plane of reference. The stream function for a single circular vortex (x_1, y_1) is given by

$$\frac{\mu}{\pi}\chi_1 = \sqrt{(xx_1)}f_1,$$

where
$$f_1 = \frac{1+k'^2}{2k}\,F - \frac{1}{k}\,E, \tag{1}$$

in which
$$k^2 = \frac{4x_1x}{(x+x_1)^2+(y-y_1)^2}. \tag{2}$$

Also the velocity of progression due to its own constitution is

$$U = \frac{\mu}{4\pi x_1}\left(\log_e \frac{8x_1}{a} - \frac{1}{4}\right)$$

where a is the radius of the cross-section of the filament. With varying aperture this will alter so as to keep the volume of the filament constant. If this be denoted by $m = 2\pi x\,.\,\pi a^2$

$$U = \frac{\mu}{8\pi x}\left(\log_e \frac{128\pi^2}{m}x^3 - \frac{1}{2}\right)$$

$$= \frac{3\mu}{8\pi x}\log_e \lambda x, \tag{3}$$

where
$$\lambda = \left(\frac{128\pi^2}{m\sqrt{e}}\right)^{1/3}.$$

In what follows we shall omit the π and suppose it contained in the μ. It will be noticed that, as in the case of the straight filaments, $\overline{\chi_1}|_2 = \overline{\chi_2}|_1$ and k is the same quantity as that denoted by f in the treatment of the two-dimensional case. Hence

$$\frac{dk}{dx_1} = \frac{k}{2x_1} - \frac{k^3(x_1+x_2)}{4x_1x_2}, \qquad \frac{dk}{dy_1} = -\frac{k^3(y_1-y_2)}{4x_1x_2},$$

$$\frac{dk}{dx_2} = \frac{k}{2x_2} - \frac{k^3(x_1+x_2)}{4x_1x_2}, \qquad \frac{dk}{dy_2} = \frac{k^3(y_1-y_2)}{4x_1x_2}.$$

It follows as before, writing y for $y_1 - y_2$ that

$$\left.\begin{aligned} \frac{dx_1}{dt} &= \frac{\mu_2}{x_1}\overline{\frac{d\chi_2}{dy}}\bigg|_1 = \mu_2\frac{k^3}{4x_1^2x_2}\,y\,\frac{d\chi}{dk} \\ \frac{dx_2}{dt} &= \dots \qquad\quad = -\mu_1\frac{k^3}{4x_1x_2^2}\,y\,\frac{d\chi}{dk} \end{aligned}\right\}. \tag{4}$$

whence $$\mu_1 x_1 \frac{dx_1}{dt} + \mu_2 x_2 \frac{dx_2}{dt} = 0,$$

or $$\mu_1 x_1^2 + \mu_2 x_2^2 = \mathrm{A} = \mu_1 a^2 + \mu_2 b^2, \tag{5}$$

where a, b are the radii of the apertures in the co-planar position.

We can therefore express x_1, x_2 in terms of one variable θ, where

$$x_1 = \sqrt{\frac{\mathrm{A}}{\mu_1}} \cos\theta, \qquad x_2 = \sqrt{\frac{\mathrm{A}}{\mu_2}} \sin\theta.$$

Further, write $\mu_1 = \mu\cos^2\alpha$, $\mu_2 = \mu\sin^2\alpha$ and let d denote the radius of the mean aperture of the rings, weighted by their respective circulations, that is

$$(\mu_1 + \mu_2)\, d^2 = \mu_1 a^2 + \mu_2 b^2.$$

Then $$\mathrm{A} = (\mu_1 + \mu_2)\, d^2 = \mu d^2,$$

$$x_1 = d\cos\theta/\cos\alpha, \qquad x_2 = d\sin\theta/\sin\alpha,$$

$$x_1 x_2 = d^2 \sin 2\theta/\sin 2\alpha, \qquad x = x_1 - x_2 = 2d\sin(\alpha - \theta)/\sin 2\alpha.$$

Again write $$\mu_1 x_1^2 - \mu_2 x_2^2 = \mathrm{X},$$

so that $$\mathrm{X} = \mu d^2 \cos 2\theta \text{ and}$$

$$\frac{d\mathrm{X}}{dt} = \mu_1\mu_2 \frac{k^3 y}{x_1 x_2} \frac{d\chi}{dk}. \tag{6}$$

Again

$$\frac{dy_1}{dt} = \frac{3\mu_1}{8x_1} \log\lambda_1 x_1 + \frac{\mu_2}{x_1} \overline{\frac{d\chi_2}{dx}}\bigg|_1, \qquad \frac{dy_2}{dt} = \frac{3\mu_2}{8x_2} \log\lambda_2 x_2 + \frac{\mu_1}{x_2} \overline{\frac{d\chi_1}{dx}}\bigg|_2.$$

Write

$$\phi(\mathrm{X}) = \frac{3\mu_1}{8x_1} \log\lambda_1 x_1 - \frac{3\mu_2}{8x_2} \log\lambda_2 x_2$$

$$= \frac{3\mu}{8d} \left\{ \frac{\cos^2\alpha}{\cos\theta} \log\left(\lambda_1 d \frac{\cos\theta}{\cos\alpha}\right) - \frac{\sin^2\alpha}{\sin\theta} \log\left(\lambda_2 d \frac{\sin\theta}{\sin\alpha}\right) \right\}.$$

Then

$$\frac{dy}{dt} = \phi(\mathrm{X}) + \frac{\mu_2}{x_1}\frac{d\chi}{dx_1} - \frac{\mu_1}{x_2}\frac{d\chi}{dx_2}$$

$$= \phi(\mathrm{X}) - \tfrac{1}{2} \frac{\mathrm{X}}{x_1^2 x_2^2} \chi + \left\{ -\tfrac{1}{2} \frac{\mathrm{X}}{x_1^2 x_2^2} + \frac{x_1 + x_2}{4x_1^2 x_2^2} (\mu_1 x_1 - \mu_2 x_2)\, k^2 \right\} k \frac{d\chi}{dk}.$$

Further

$$\frac{dk}{dt} = \frac{dk}{dx_1}\frac{dx_1}{dt} + \ldots$$

$$= \tfrac{1}{2} k \left(\frac{1}{x_1}\frac{dx_1}{dt} + \frac{1}{x_2}\frac{dx_2}{dt} \right) - \frac{k^3(x_1 + x_2)}{4x_1 x_2} \left(\frac{dx_1}{dt} + \frac{dx_2}{dt} \right) - \frac{k^3 y}{4x_1 x_2} \frac{dy}{dt}.$$

Substituting the values for dx/dt, dy/dt given above, it can be shown that

$$\frac{dk}{dt} = \left\{ \frac{\mathrm{X}}{8x_1^3 x_2^3} \chi - \frac{1}{4x_1 x_2} \phi(\mathrm{X}) \right\} k^3 y.$$

Substituting for k^3y/x_1x_2 from (6)

$$\frac{d\chi}{dk}\frac{dk}{dX}=\frac{1}{4\mu_1\mu_2}\left\{\frac{X}{2x_1^2x_2^2}\chi-\phi(X)\right\},$$

$$\sqrt{(x_1x_2)}\frac{df}{dX}=\frac{1}{4\mu_1\mu_2}\left\{\frac{X}{2(x_1x_2)^{3/2}}f-\phi(X)\right\},$$

or transforming to θ

$$\sqrt{\left(\frac{\sin 2\theta}{\sin 2\alpha}\right)}\frac{df}{d\theta}+\frac{\cos 2\theta}{\sqrt{(\sin 2\alpha\sin 2\theta)}}f=\frac{2d}{\mu}\frac{\sin 2\theta}{\sin^2 2\alpha}\phi(\theta),$$

$$\frac{d}{d\theta}(f\sqrt{\sin 2\theta})=\frac{3}{8\sqrt{2}}\left\{\cot^{3/2}\alpha\sin\theta\log\lambda_1 d\frac{\cos\theta}{\cos\alpha}\right.$$

$$\left.-\tan^{3/2}\alpha\cos\theta\log\lambda_2 d\frac{\sin\theta}{\sin\alpha}\right\}.$$

Hence

$$f\sqrt{\sin 2\theta}=C-\frac{3}{8\sqrt{2}}\{\cot^{3/2}\alpha\cos\theta\log_e(P_1\cos\theta)$$

$$+\tan^{3/2}\alpha\sin\theta\log_e(P_2\sin\theta)\}$$

$$=C-0{\cdot}61108\,\{\cot^{3/2}\alpha\cos\theta\log_{10}(P_1\cos\theta)+\ldots\},\qquad(7)$$

where $P_1=\lambda_1 d/(e\cos\alpha)$, $P_2=\lambda_2 d/(e\sin\alpha)$.

If c denotes the radius of a sphere whose volume is that of a ring

$$\lambda=\left(\frac{128\pi^2}{m\sqrt{e}}\right)^{1/3}=\frac{1}{c}\left(\frac{96\pi}{\sqrt{e}}\right)^{1/3}$$

and $$P_1=\frac{2{\cdot}9533}{\sqrt{2}\cos\alpha}\frac{d}{c_1},\qquad P_2=\frac{2{\cdot}9533}{\sqrt{2}\sin\alpha}\frac{d}{c_2}.\qquad(7a)$$

When the rings have equal circulations,

$$P_1=P_2=2{\cdot}9533\,d/c.$$

In all cases where our present supposition of thin filaments holds, d is always a very large multiple of c, of the order 100 or larger. It will be convenient to suppose it given by 10^n. It corresponds to a cross-section whose radius $= 2{\cdot}122\times 10^{-(3n+1)/2}d$.

The equation to the relative path may be written in any given case

$$f\sqrt{\sin 2\theta}=C-L\cos(\theta-\gamma)-0{\cdot}61108\cot^{3/2}\alpha\cos\theta\log\cos\theta$$

$$-0{\cdot}61108\tan^{3/2}\alpha\sin\theta\log\sin\theta.\qquad(8)$$

The values of the functions $0{\cdot}61108\cos\theta\log_{10}\cos\theta$, $0{\cdot}61108\sin\theta\log_{10}\sin\theta$ are given at intervals of 5° from 0 to 45 in Table I. It is not necessary to give them for higher values since for $\pi/2-\theta$ they interchange their values. The values of f are given in Table II. In calculating any path, then, f should first be represented by a curve drawn on squared paper with abscissæ $= \phi$, where $k = \sin\phi$. This should be done once for all, and serves for the determination of any path. If the right-hand side of the equation is divided by $\sqrt{\sin 2\theta}$, the path is determined by an equation of the form

$$f(\phi) = \mathbf{f}(\theta).$$

Then $\mathbf{f}(\theta)$ is calculated for any θ. The f curve then gives the value of ϕ for which $f(\phi)$ has this value. Then x_1, x_2, $x = x_1 - x_2$ are given by θ, and y is found from $k = \sin\phi$ by

$$k^2 = \frac{4x_1x_2}{(x_1+x_2)^2+y^2},$$

or
$$y = \frac{2d}{\sin 2\alpha}\left\{\frac{\sin 2\theta \sin 2\alpha}{k^2} - \sin^2(\theta+\alpha)\right\}^{1/2}$$

and
$$x = \frac{2d}{\sin 2\alpha}\sin(\alpha-\theta).$$

The foregoing discussion is completely general. In the application, however, to actual cases a modification is required when the two circulations are oppositely directed. We take first the case when these are the same.

8. *Conditions for Permanent Union* (μ_1, μ_2 same sign).—We may suppose any given arrangement of the rings defined by their apertures at the instant when the rings are co-planar. If θ_0 be the value of θ at this instant

$$b/a = \tan\theta_0 \cot\alpha.$$

At this instant, $y = 0$ and

$$\cos\phi = k' = \frac{a-b}{a+b},$$

$$\therefore \qquad \frac{b}{a} = \frac{1-\cos\phi}{1+\cos\phi} = \tan^2\tfrac{1}{2}\phi = \tan\theta_0\cot\alpha, \qquad (9)$$

so that the value of f (say f_0) at this instant is determinable in terms of θ_0. Putting in these values in equation (7) determines the value of the constant C for the given arrangement.

There will be some critical value of θ_0 below which one ring passes through the other and separates to infinity, and above which the two will remain in permanent union, continually threading each other. To determine this, consider the arrangement in the critical case when, say, ring II was at an infinite distance before reaching I. The two rings must now have equal velocities of progression. If that of II were just less it would never reach I. If it were just greater it would ultimately reach I, and since the paths are symmetrical, pass off to infinity at the other side. Thus the critical state (θ_0) corresponds to one in which the two rings have equal velocities of progression when they have separated so far as to be beyond each other's direct influence.

If their radii are then x_1, x_2, this condition gives

$$\frac{\mu_1}{x_1} \log \lambda_1 x_1 = \frac{\mu_2}{x_2} \log \lambda_2 x_2,$$

or

$$\tan\theta \, \frac{\log (\mathrm{P}_1 e \cos\theta)}{\log (\mathrm{P}_2 e \sin\theta)} = \tan^3 \alpha. \qquad (10)$$

Let θ_1 be the solution of this equation. It must also satisfy the equation to the path. Now, at infinity $k = 0$ and $f = 0$. Hence

$$\mathrm{C} = \mathrm{L} \cos(\theta_1 - \gamma) + 0{\cdot}61108 \cot^{3/2} \alpha \cos\theta_1 \log\cos\theta_1 + \ldots .$$

The critical value θ_0 is therefore the root of the equation

$$f_0 \sqrt{\sin 2\theta_0} = \mathrm{L}\cos(\theta_1 - \gamma) + 0{\cdot}61108 \cot^{3/2} \alpha \cos\theta_1 \log\cos\theta_1 + \ldots$$
$$- \mathrm{L}\cos(\theta_0 - \gamma) - 0{\cdot}61108 \cot^{3/2} \alpha \cos\theta_0 \log\cos\theta_0 - \ldots . \quad (11)$$

When the circulations are equal, the solution of (10) is $\theta_1 = \frac{1}{4}\pi$, as it clearly should be since then $b = a$.

9. *Case of Equal Rings.*—Here $\mu_1 = \mu_2$, $\alpha = \pi/4$, $\gamma = \pi/4$, and $c_1 = c_2$. Also $\theta_1 = \pi/4$. The critical equation is now

$$f_0 \sqrt{\sin 2\theta_0} = 0{\cdot}86419 n + 0{\cdot}27636 - (0{\cdot}86419 n + 0{\cdot}40603) \cos(\pi/4 - \theta_0)$$
$$- 0{\cdot}61108 (\cos\theta_0 \log\cos\theta_0 + \sin\theta_0 \log\sin\theta_0), \quad (11a)$$

f_0 is calculated as a function of θ_0 since $\tan^2 \frac{1}{2}\phi = \tan\theta_0$. The curves are drawn for $\eta =$ left-hand side and $\eta =$ right-hand side. It is only necessary to do this for $\theta < 45°$, since now the curves are the same for θ and $\pi/2 - \theta$. The first curve increases continuously from 0 to ∞, the second decreases from a finite positive value to zero. Hence they must intersect at a single point which uniquely determines θ_0. For the case of $n = 3$, or $d = 1000\,c$, this root is given by $\theta = 23° \; 30'$. Hence the critical ratio of the apertures is given by

$$b/a = \tan 23° \; 30' = 0{\cdot}4348.$$

With equal rings in which $d = 1000\,c$ then the ratio of the apertures, when co-planar, must be greater than 0·4348, if they are to remain in permanent union.

The above analysis may be illustrated by considering two arrangements, one near the critical state, the other considerably greater, when $d = 1000\,c$. The relative paths for $b/a = \frac{1}{2}$ and $\frac{3}{4}$ or $\theta_0 = 26° \; 34'$ and $36° \; 52'$ are represented by curves A and B in fig. 5.

The foregoing treatment of equation (11a), to find the critical ratio, when d/c is given, is, however, too restricted. The equation shows that any value of the critical ratio is possible provided the absolute values of the apertures, or of d, are properly chosen. In fact, the equation may be written

$$n+0\cdot4698 = 1\cdot1571\frac{f\sqrt{\sin 2\theta_0}+0\cdot1300+0\cdot61108(\cos\theta_0\log\cos\theta_0+\dots)}{1-\cos(\pi/4-\theta_0)}.$$

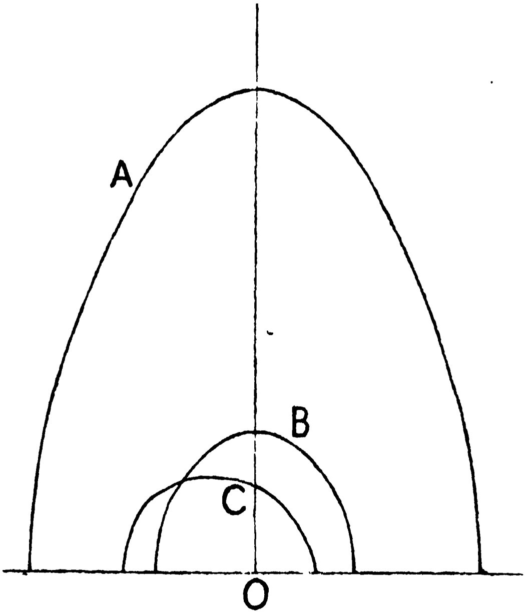


Fig. 5.

The numerator of the fraction on the right is always positive, since the maximum of the negative terms in logs is 0·130 when $\theta_0 = \pi/4$. As θ_0 increases from 0 to $\pi/4$, the denominator decreases from $1-1/\sqrt{2}$ to 0, whilst the $f\sqrt{\sin 2\theta_0}$ increases from 0 to ∞. Hence $n+0\cdot4698$ increases from $1\cdot1571\times 0\cdot1300(2+\sqrt{2}) = 0\cdot5138$ to ∞, or n from 0·44 to ∞. But in order that our stream function may be valid, not only must d/c be large, but the smaller aperture, $b = d\sqrt{2}\sin\theta_0$, must be large compared with the radius of its cross-section. As an extreme case for the formula validity, take this ratio > 10, which corresponds to $b > 10^{0\cdot551}c$, whence

$$d\sqrt{2}\sin\theta_0 > 10^{0\cdot551}c,$$

$$n > 0\cdot551 + \log\operatorname{cosec}\theta_0 - 0\cdot150 > 0\cdot400 + \log\operatorname{cosec}\theta_0.$$

From the equation we find

θ_0.	n.
10°	0·0548
15°	0·4456
20°	1·432
25°	4·211

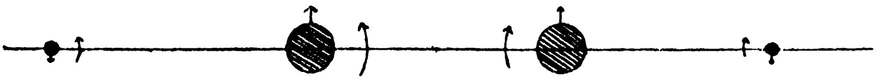
Hence n is too small for 10°, 15°. For $\theta = 20°, 0{\cdot}400 + \log \operatorname{cosec} 20° = 0{\cdot}867$, so that for this angle, and above, the value of n comes within the limits of formula validity. The limit is found close to 18°, for which $n = {\cdot}910$ and $0{\cdot}400 + \log \operatorname{cosec} 18° = 0{\cdot}917$. This gives

$$b/a = 0{\cdot}325; \quad d/c = 10^{0{\cdot}910}; \quad b/a = 10^{0{\cdot}360}; \quad a/c = 10^{1{\cdot}038},$$

whence if r denote the radius of cross-section of the filament

$$a/r_1 = 52{\cdot}6, \qquad b/r_2 = 5{\cdot}17.$$

This system is drawn to scale in the following diagram—



The configuration is clearly one in which the assumption of a thin filament does not hold, and the stream function is no longer approximate. But it is sufficiently close to show that the least possible ratio of b/a for permanent union is close to tan 19°, or (say) one-third, with a mean aperture given by $d = 10c$. If a less ratio is possible, it must be accompanied by considerable deformations of the cross-section of the inner, which also must have considerable thickness.

If the ratio b/a for the limiting case approaches unity, n is exceedingly large. Are the calculations now invalid, by a too close approximation of the two rings? To answer this, take $\theta_0 = \pi/4 - \theta$, with θ tending to zero. Then

$$a = d\sqrt{2}\cos(\pi/4-\theta) = d(1+\theta),$$

$$b = d\sqrt{2}\sin(\pi/4-\theta) = d(1-\theta),$$

$$a-b = 2d\theta,$$

also the cross-sectional radii will tend to equality, say r, where

$$r = 2{\cdot}122 \times 10^{-(3n+1)/2}\, d,$$

$$\therefore \qquad (a-b)/r = \theta\, 10^{(3n+1)/2},$$

also n is now of the order

$$n = \frac{f}{\frac{1}{2}\theta^2} = \frac{2f}{\theta^2} = \frac{1}{\theta^2}\log\sec\phi$$

and
$$\tan\frac{\phi}{2} = \sqrt{\tan\left(\frac{\pi}{4}-\theta\right)} = 1-\theta.$$

Hence
$$\phi = \frac{\pi}{2}-\theta \quad \text{and} \quad n = \frac{1}{\theta^2}\log\frac{1}{\theta}.$$

Thus
$$\frac{a-b}{r} \text{ is of order } \theta\left(\frac{1}{\theta}\right)^{3/\theta^2},$$

or the distance of the rings from one another is infinitely larger than the cross-sections of the filaments. Consequently the above results are valid when a, b approach equality.

10. *Case of Different Circulations.*—In choosing illustrations for this kind of motion, the volumes of the rings, as well as their circulations, have to be taken account of. We shall for this purpose confine our attention to configurations where the vortical matter is of the same quality, or the vorticity is the same. In this case, if ω denote the vortical rotation at any point in a ring at a distance x from the straight axis, $\omega = kx$, where k is a constant. Hence

$$\mu = 2\pi\omega r^2 = 2\pi kxa^2 = \frac{k}{\pi} m = \frac{4}{3} kc^3.$$

Thus $$\cot^2 \alpha = \mu_1/\mu_2 = m_1/m_2 = (c_1/c_2)^3.$$

Take c to represent the radius of that sphere whose diametral area is the sum of those of the rings, so that

$$c^2 = c_1^2 + c_2^2$$

and put $$c_1 = c \cos\beta, \qquad c_2 = c \sin\beta.$$

Then $\cot^2 \alpha = \cot^3 \beta$, and equations (7$a$) become

$$P_1 = \frac{2{\cdot}9533}{\sqrt{2 \cos\alpha \cos\beta}} \frac{d}{c}, \qquad P_2 = \frac{2{\cdot}9533}{\sqrt{2 \sin\alpha \sin\beta}} \frac{d}{c}.$$

For a case in illustration take $d = 1000\,c$ and $\cot^{3/2} \alpha = 3$. This gives

$$\alpha = 25^\circ\ 40'\ 34'', \qquad \beta = 31^\circ\ 32'\ 11'',$$
$$\mu_1/\mu_2 = \cot^2 \alpha = 4{\cdot}3267,$$
$$c_1/c_2 = 1{\cdot}6295,$$
$$\log P_1 = 3{\cdot}4343, \qquad \log P_2 = 3{\cdot}9645.$$

In order to determine the condition of permanent union, equation (10) is found to be

$$\log \left\{ \tan\theta \, \frac{3{\cdot}8686 - \log \sec\theta}{4{\cdot}3988 - \log \operatorname{cosec}\theta} \right\} = 1{\cdot}04576.$$

The root is $\theta_1 = 5^\circ\ 33'\ 18''$. This makes the critical constant $C = 6{\cdot}3197$ and the equation to determine the critical θ_0 is

$$f\sqrt{\sin 2\theta} = 6{\cdot}3197 - 6{\cdot}3474 \cos(\theta - 7^\circ\ 19') - 3\,F(\theta) - \tfrac{1}{3}\,F(\pi/2 - \theta)$$

where $$F(\theta) \equiv 0{\cdot}61108 \cos\theta \log_{10} \cos\theta,$$

and f is a function of θ given by

$$\tan^2 \tfrac{1}{2}\phi = 3^{2/3} \tan\theta.$$

There will be two roots, one corresponding to the configuration where the weaker ring is inside, the other to that in which it is outside. They are respectively very close to $\theta_0 = 2^\circ\,30'$ and $\theta_0 = 32^\circ$, corresponding to

$$b/a = 0{\cdot}037 \text{ (weaker inside)},$$

$$b/a = 0{\cdot}769 \text{ (weaker outside)}.$$

Thus if the weaker is placed by chance inside the stronger it is practically certain to remain in union. If, however, the stronger is put by chance inside the weaker the chances are three to one against the partnership being permanent.

The curve C in fig. 5, represents the relative path of the weaker as seen from the stronger, when $\theta_0 = 15^\circ$ for weaker inside. It corresponds to ratio of radii of weak to strong of 0·557. For this case the value of C in the equation to the path is 6·4974.

11. *Opposite Circulations.*—With $\mu_2 = -\mu_2$ equation (5) becomes

$$\mu_1 x_1^2 - \mu_2 x_2^2 = \text{A}, \tag{12}$$

with

$$\mu_1 x_1^2 + \mu_2 x_2^2 = \text{X}.$$

The former transformations may be replaced by*

$$x_1 = \sqrt{\frac{\text{A}}{\mu_1}}\cosh\theta, \qquad x_2 = \sqrt{\frac{\text{A}}{\mu_2}}\sinh\theta,$$

$$\mu_1 = \mu\cosh^2\alpha, \qquad \mu_2 = \mu\sinh^2\alpha,$$

$$(\mu_1 - \mu_2)\,d^2 = \mu d^2 = \text{A},$$

so that

$$x_1 = d\cosh\theta/\cosh\alpha, \qquad x_2 = d\sinh\theta/\sinh\alpha,$$

$$x = 2d\sinh(\alpha - \theta)/\sinh 2\alpha, \qquad \text{X} = \mu d^2\cosh 2\theta,$$

also when co-planar, the value of ϕ in f is given by

$$\tan^2\tfrac{1}{2}\phi = \coth\alpha\tanh\theta.$$

In fact, the formulæ obtained for μ_1, μ_2 positive are transformed by writing θi, αi.

In the case of like circulations, with $\mu_1 x_1^2 + \mu_2 x_2^2$ constant, it is clear that the apertures can never vary largely from those in the standard co-planar position. Here, however, any values up to infinity may enter so far as condition (12) is concerned. If the circulations are equal it is clear that $x = x_1 - x_2$ can never change sign, that is, one ring can never cross the path of the other. When the circulations are not equal, this crossing is possible but only when the two apertures have opened out to extremely large amounts, unless the two μ_1, μ_2 differ considerably. With equal μ, the θ, α, μ, etc., become indeterminate, and it is better to treat this case independently.

---

* As existing tables are more complete for circular than for hyperbolic functions, it might be advisable to use the unsymmetrical transformation $\mu_1 = \mu\sec^2\alpha$, $\mu_2 = \mu\tan^2\alpha$, $x_1 = d\sec\theta/\sec\alpha$, $x_2 = d\tan\theta/\tan\alpha$, $\text{X} = \mu d^2(1 + \sin^2\theta)/\cos^2\theta$, $x = 2d\cos\alpha\sec\theta(1 - \sin\theta\operatorname{cosec}\alpha)$, $\tan^2\frac{1}{2}\phi = \sin\theta/\sin\alpha$.

12. *Conditions for Union with Opposite Circulations.*—Consider the critical co-planar state which is the boundary between configurations of permanent and temporary union. If the inner ring is slightly smaller it will pass through and on to infinity. If it is larger, its own velocity of progression is smaller than that due to the other ring and it is swept back, and its aperture still further increased by the other. In the critical condition, therefore, the two rings must just be at relative rest or $dy/dt = 0$ when $y = 0$. Hence

$$\phi\,(\mathrm{X}) - \tfrac{1}{2}\frac{\mathrm{X}}{x_1^2 x_2^2}\chi + \left\{-\tfrac{1}{2}\frac{\mathrm{X}}{x_1^2 x_2^2} + \frac{x_1+x_2}{4x_1^2 x_2^2}(\mu_1 x_1 + \mu_2 x_2)k^2\right\} k\frac{d\chi}{dk} = 0,$$

whilst the simultaneous condition $y = 0$ gives $k = \sin\phi$ from

$$\tan^2 \tfrac{1}{2}\phi = x_2/x_1.$$

In any given pair of rings μ_1/μ_2 and other constants are known, and this is an equation to determine x_1/x_2 in the critical configuration. Our present purpose, however, is to obtain information as to the general nature of the motions. It will be simpler, therefore, to take an instance of x_1, x_2 given and determine the values of $\mu_1/\mu_2 (\equiv z \text{ say})$, which will make this configuration a critical one. We will take

$$x_1 = nd, \qquad x_2 = d, \qquad \tan\tfrac{1}{2}\phi = n^{-1/2}, \qquad (n > 1).$$

The equation of condition may now be written

$$\mu_1\left\{nf + n\left(1 - \frac{n+1}{2n}k^2\right)k\frac{df}{dk} - \frac{3}{4\sqrt{n}}\log_e \lambda_1 x_1\right\}$$
$$= \mu_2\left\{\frac{3\sqrt{n}}{4}\log\lambda_2 x_2 - f - \left(1 - \frac{n+1}{2n}k^2\right)k\frac{df}{dk}\right\}.$$

Here $\lambda_1 = \left(\dfrac{96\pi}{\sqrt{e}}\right)^{1/3}\dfrac{1}{c_1}, \qquad \lambda_2 = \left(\dfrac{96\pi}{\sqrt{e}}\right)^{1/3}\dfrac{1}{c_2}.$

Further if the vorticities of the two rings are the same

$$\mu_1/\mu_2 = (c_1/c_2)^3 = z,$$

$$\therefore \qquad \lambda_1 x_1 = \lambda_2 z^{-1/3} n \,.\, x_2 = \left(\frac{96\pi}{\sqrt{e}}\right)^{1/3}\frac{n}{z^{1/3}}\cdot\frac{d}{c_2}.$$

Write $\left(\dfrac{96\pi}{\sqrt{e}}\right)^{1/3} = 5{\cdot}676 = \lambda, \qquad f + \left(1 - \dfrac{n+1}{2n}k^2\right)k\dfrac{df}{dk} = \gamma.$

Then

$$z = n\,\frac{\log_{10}\lambda\, d/c - 0{\cdot}57905\,\gamma/\sqrt{n}}{0{\cdot}57905\, n\sqrt{n}\gamma - \log n + \frac{1}{3}\log z - \log_{10}\lambda\, d/c}.$$

In this

$$\gamma = f + \left(1 - \frac{n+1}{2n}k^2\right)k\frac{df}{dk} \text{ and}$$

$$k\frac{df}{dk} = -\frac{1}{k}\,\mathrm{F} + \frac{1+k'^2}{kk'^2}\,\mathrm{E}.$$

In general it is clear that for a positive value of z it is necessary that

$$\log \lambda_2 x_2 > \frac{4}{3\sqrt{n}}\gamma, \qquad > \frac{1}{n^2}\log \lambda_1 x_1,$$

or

$$(\lambda_2 x_2)^{n^2} > \lambda_1 x_1 \qquad > n z^{-1/3} \lambda_2 x_2,$$

$$\left(\lambda \frac{d}{c}\right)^{n^2-1} > n z^{-1/3}$$

For illustrative cases take $x_2/x_1 = 0{\cdot}81$ and $0{\cdot}5$.

$$x_2/x_1 = 0{\cdot}81, \qquad \tan \tfrac{1}{2}\phi = 0{\cdot}9, \qquad \phi = 84^\circ.$$

This gives $\gamma = 10{\cdot}715$ and

$$z = \frac{100}{81}\,\frac{\log \lambda\, d/c - 5{\cdot}584}{8{\cdot}475 + \frac{1}{3}\log z - \log \lambda\, d/c}.$$

Hence $$\lambda\, d/c > 10^{6{\cdot}6} \quad < 10^{8{\cdot}4}.$$

If $$\frac{\lambda d}{c} = 10^6, \qquad z = 1{\cdot}234\,\frac{0{\cdot}416}{2{\cdot}475 + \frac{1}{3}\log z} = 0{\cdot}22.$$

If $$\frac{\lambda d}{c} = 10^8, \qquad z = 1{\cdot}234\,\frac{2{\cdot}416}{0{\cdot}475 + \frac{1}{3}\log z} = 4{\cdot}3.$$

For the first the stronger is inside, for the second outside. For all possible cases the filament will be very thin, of order $10^{-7}\times$ the aperture.

$$x_2/x_1 = 0{\cdot}5, \qquad \phi = 70^\circ\, 32', \qquad \gamma = 4{\cdot}155,$$

$$z = 2\,\frac{\log \lambda\, d/c - 1{\cdot}701}{6{\cdot}504 + \frac{1}{3}\log z - \log \lambda\, d/c}$$

$$= 2\,\frac{\log d/c - 0{\cdot}947}{5{\cdot}750 + \frac{1}{3}\log z - \log d/c}.$$

For comparison with the cases considered with like circulations take $\log d/c = 3$,

$$z = 2\,\frac{2{\cdot}053}{2{\cdot}750 + \frac{1}{3}\log z} = 1{\cdot}46.$$

13. *Equal and Opposite Rings.*—In this case as we have seen one ring can never cross the path of the other, that is the ring with the smaller aperture at any instant will always have the smaller aperture. We consider first the configuration of equal apertures, when the apertures are then always equal. It is the case indeed in which a single ring moves towards a plane boundary which is parallel to its own plane. Refer the motion to the mid-plane (or the fixed boundary). Then the co-ordinates of the ring are $x = x_1 = x_2$ and $y = y_1 = -y_2$, and

$$k^2 = \sin^2\phi = \frac{x^2}{x^2+y^2}, \qquad \tan\phi = x/y.$$

Hence if O be the point where the central axis cuts the plane, and P the centre of the cross-section of the filament, $OP = r$ and OP makes an angle ϕ with the direction of translation. In fact r, ϕ are the polar co-ordinates of the position of the ring.

The equation to the path is given by

$$x\frac{df}{dX} = \frac{3}{16\mu x}\log \lambda x - \frac{1}{4\pi x}f,$$

or since $X = 2\mu x^2$

$$d/dx\,(xf) = \tfrac{3}{4}\log \lambda x,$$

whence

$$xf = l + \tfrac{3}{4}\,x\log \lambda x/e,$$

where l is a constant of integration. Also $x = r\sin\phi = rk$,

$$\therefore \qquad \frac{l}{r} = -\tfrac{3}{4}\,k\log\frac{\lambda rk}{e} + (1-\tfrac{1}{2}\,k^2)\,F - E.$$

When ϕ approaches $\frac{1}{2}\pi$ or k' approaches zero

$$\frac{l}{r} = -\tfrac{3}{4}\log\frac{\lambda r}{c} - 1 + \tfrac{1}{2}\log\frac{4}{k'}$$

$$= \tfrac{1}{4}\log\frac{16}{\lambda^3 e}\cdot\frac{1}{k'^2 r^3}.$$

Hence r becomes infinite at the rate

$$r^3 = \frac{16}{\lambda^3 e k'^2},$$

or

$$r = \frac{2}{\lambda}\left(\frac{2}{e}\right)^{1/3}\sec^{2/3}\phi = A\sec^{2/3}\phi \text{ as } \phi \text{ tends to } \tfrac{1}{2}\pi,$$

also $y = r\cos\phi = A\cos^{1/3}\phi$.

If c is the radius of the cross-section at any time

$$rc^2 = \text{const.} = r_0 c_0^2,$$

$$\frac{c}{y} = \frac{c_0 r_0^{1/2}}{A^{3/2}} = \text{constant}.$$

If then c/y is sufficiently small at any time, *i.e.* the filament sufficiently thin, to justify our use of the formula for χ, the formula will hold even when the two rings become infinitely close when their apertures are infinitely large, and the equation to the path will remain throughout that given above, if no other conditions occur. But if the filaments decrease indefinitely a state will intervene, depending on the pressure of the fluid, at which a hollow is formed* and the limiting case is one of constant cross-section. Unless, therefore, the pressure is infinitely large some secondary changes will intervene after a certain point, the cross-sections will not be circular, and the configuration altered in some way not hitherto determined.

---

* 'Proc. Camb. Phil. Soc.,' vol. 3, p. 284 (1879).

14. *Path in the fluid.*—Write $y_1+y_2=\mathrm{Y}$. Then

$$\frac{d\mathrm{Y}}{dt}=\psi(\mathrm{X})+\frac{1}{2}\frac{\mathrm{A}}{x_1^2x_2^2}\chi+\left\{\frac{1}{2}\frac{\mathrm{A}}{x_1^2x_2^2}-\frac{(x_1+x_2)}{4x_1^2x_2^2}(\mu_1x_1+\mu_2x_2)k^2\right\}k\frac{d\chi}{dk}$$

$$=\psi(\mathrm{X})+\frac{1}{2}\frac{\mathrm{A}}{x_1^2x_2^2}\chi+\left\{\frac{1}{2}\frac{\mathrm{A}}{x_1^2x_2^2}-\frac{\mathrm{A}+(\mu_1+\mu_2)x_1x_2}{4x_1^2x_2^2}k^2\right\}k\frac{d\chi}{dk},$$

where

$$\psi(\mathrm{X})=\frac{3\mu_1}{8x_1}\log\lambda_1x_1+\frac{3\mu_2}{8x_2}\log\lambda_2x_2.$$

Also (6)

$$\frac{d\mathrm{X}}{dt}=\mu_1\mu_2\frac{k^3y}{x_1x_2}\frac{d\chi}{dk},$$

$$\therefore\quad\frac{d\mathrm{Y}}{d\mathrm{X}}=\frac{\psi(\mathrm{X})}{\mu_1\mu_2k^3}\frac{\sqrt{(x_1x_2)}}{y}\cdot\frac{1}{df/dk}+\frac{\mathrm{A}}{2\mu_1\mu_2x_1x_2y}\cdot\frac{f}{df/dk}$$

$$+\frac{\mathrm{A}}{2\mu_1\mu_2x_1x_2y}\left\{\frac{1}{k^2}-\frac{1+(\mu_1+\mu_2)x_1x_2}{2\mathrm{A}}\right\}.$$

It is, of course, not integrable directly, but in any special case in tracing the relative path, contemporaneous values of k, y, and X (or θ) are known. The curve representing the right hand in terms of X as abscissa can consequently be traced, and Y obtained in terms of X by mechanical quadratures.

15. *The Time.*—The relation to the time again cannot be found in general terms, but it may be found for special cases by (6)

or

$$\frac{dt}{d\mathrm{X}}=\frac{\sqrt{(x_1x_2)}}{\mu_1\mu_2k^3y}\frac{1}{df/dk}$$

in the same way as Y in the previous article.

Table I.

θ.	$0\cdot61108\cos\theta\log_{10}\cos\theta$.	$0\cdot61108\sin\theta\log_{10}\sin\theta$.
°		
5	−0·00100	−0·05631
10	−0·00400	−0·08068
15	−0·00883	−0·09284
20	−0·01551	−0·09738
25	−0·02366	−0·09660
30	−0·03306	−0·09198
35	−0·04336	−0·08461
40	−0·05418	−0·07589
45	−0·06504	−0·06504

Table II.

$f = (F - E)/k - \frac{1}{2}kF, \qquad k = \sin\phi.$

ϕ.	f.	ϕ.	f.
0°	0·	55	0·11398
10	0·00049	60	0·1579
15	0·00180	65	0·2172
20	0·00426	70	0·2982
30	0·01515	75	0·4146
35	0·02461	80	0·5932
40	0·03693	85	0·9213
45	0·05647	89	1·7199
50	0·08106	90	∞

A Spectroscopic Investigation of the Ionisation of Argon by Electron Collisions.

By FRANK HORTON, Sc.D., Professor of Physics in the University of London, and ANN CATHERINE DAVIES, D.Sc., Royal Holloway College.

(Communicated by Sir J. J. Thomson, O.M., F.R.S. Received June 17, 1922.)

Introduction.

The authors have previously described an investigation of the effects of electron collisions with argon atoms,* in which the production of radiation and of ionisation was looked for, using a delicate electrometer as the detecting instrument. As the result of these experiments, it was found that the value of the minimum electron energy for the production of resonance radiation from argon was 11·5 volts, and the minimum electron energy necessary for the detachment of a single electron from a normal argon atom was 15·1 volts. Since the publication of these results some other investigations of the critical voltages for electrons in argon have been published. Several of these were made using argon as a gas filling for a thermionic valve, but the precise methods of arranging the electric fields, and of carrying out the experiment differed somewhat in different cases. The values of the ionisation voltage deduced by the various observers are not in good agreement. For instance, from experiments with three-electrode valves, Stead and Gossling† concluded that this value is 12·5 volts, Hodgson and Palmer‡ that it is 16·6 volts, and

---

* F. Horton and A. C. Davies, 'Roy. Soc. Proc.,' A, vol. 97, p. 1 (1920).

† G. Stead and B. S. Gossling, 'Phil. Mag.,' vol. 40, p. 413 (1920).

‡ B. Hodgson and L. S. Palmer, 'Rad. Rev.,' vol. 1, No. 11 (1920).

Déjardin[*] that it is 15·0 volts. Found,[†] using a two-electrode tube, obtained the value 15·6 volts, whereas Rentschler,[‡] using a three-electrode apparatus, had previously given 17 volts as the ionisation voltage.

In attempting to account for these discordant values it may be mentioned that in the authors' experiments, already referred to, a very marked discontinuous increase of current occurred a few volts higher than the stage at which ionisation was first detected. This increase of current took place at a potential difference which varied over a range of several volts, but which was never below the ionisation voltage. At certain pressures it was accompanied by the appearance of luminosity in the gas. The effect was attributed to the largely increased electron emission from the glowing filament when neutralisation of the space charge of the electrons in its neighbourhood occurred. It seems not improbable that the effect obtained by Rentschler at 17 volts, and interpreted by him as the beginning of ionisation, in reality indicates the stage when neutralisation of the space charge of the emitted electrons takes place. It also seems probable that the very low value (12·5 volts) obtained by Stead and Gossling must have been due to the presence in the apparatus of some impurity which was ionisable by the argon radiation.

In view of the fact that argon can be stimulated so as to give rise to two quite different spectra—the red spectrum and the blue spectrum—and that it has been concluded, in the first place by Stark and Kirschbaum,[§] and subsequently by Friedersdorff,[||] from experiments on canal rays in argon, that the emitters of the spectra are respectively singly and multiply charged argon atoms, it was anticipated that an investigation of the minimum electron energies associated with the excitation of lines belonging to the red spectrum, and lines belonging to the blue spectrum, would supply evidence as to the ionisation voltages of argon.

The advantage of the spectroscopic method of testing for ionisation, over the method of collecting positive ions, is that it gives evidence as to the nature of the positive ions produced in any particular case. Thus, although the test for the presence of positive ions can be made more delicate than the spectroscopic test, particularly in the case of the ionisation of the normal atom, and although accurate values of critical points can best be obtained from current-voltage curves, the spectroscopic test provides the more convincing evidence of the exact significance of particular critical points.

---

[*] G. Déjardin, 'Comptes Rendus,' vol. 172, p. 1347 (1921).

[†] C. Found, 'Phys. Rev.,' vol. 16, p. 41 (1920).

[‡] H. C. Rentschler, 'Phys. Rev.,' vol. 14, p. 503 (1919).

[§] J. Stark and H. Kirschbaum, 'Ann. der Phys.,' vol. 42, p. 255 (1913).

[||] K. Friedersdorff, 'Ann. der Phys.,' vol. 47, p. 737 (1915).

An investigation of the excitation of the spectra of argon by slowly moving electrons has recently been published by Déjardin,* who took a series of photographs of the spectrum of the glow obtained at different voltages in argon at a pressure of 0·1 mm. His results confirmed in general the conclusions of Stark as to the nature of the atoms emitting the different spectra, but observations were not taken at sufficiently close voltage intervals to give very precise information as to the critical electron energies.

A combination of the spectroscopic and current-voltage tests was applied by the authors to the investigation of the ionisation of neon and yielded interesting results, which have been given elsewhere.† A somewhat similar combination of the two methods was therefore employed in the case of argon, the currents between the electrodes being measured simultaneously with the taking of visual spectroscopic observations.

The experiments to be described in the present paper confirm the values 11·5 volts and 15·1 volts previously obtained by the authors as the minimum resonance radiation voltage and the minimum ionisation voltage respectively, for electrons in argon, and they show that the lines of the red spectrum are those which result from recombination of electrons and positive ions after the occurrence of simple ionisation of the argon atoms. In addition, the experiments show that certain of the "blue spectrum" lines are associated with the removal of two electrons from the normal atom, and that the minimum voltage at which they can be produced by single electron impacts is 34 volts. Under conditions which facilitated the further ionisation of already ionised atoms, it was found that the "blue spectrum" lines could be produced by 19 volts collisions, and that, with very intense bombarding electron streams, lines of this spectrum could be excited at still lower voltages, though no other definite limit associated with their excitation was obtained.

Description of Apparatus.

Two forms of apparatus were used in this investigation; one of these is represented diagrammatically in fig. 1. It resembles that used by the authors in their experiments with neon, its main features being that the spectrum of the glow can be observed in a part of the tube where the bombarding electrons suffer no change of velocity except that which results from collisions with gas atoms, and that the form of the apparatus enables it to be placed between the poles of a strong electromagnet, whereby a concentration of the luminosity into a bright column parallel to the slit of the spectroscope can be obtained. The apparatus contains two parallel platinum filaments, F, F, coated with a mixture of lime and baryta. Only one of these was used at a time. They were situated about 3 mm. above the circular piece of very fine platinum gauze, G. The piece of similar gauze, H, was 1 cm. below this, and it had a cylindrical piece of platinum gauze fixed round its edges, as seen in the figure. The anode A was a circular platinum plate about 1 cm. in diameter. The spectroscope was arranged so as to view the centre of the tube between G and H in a direction at right angles to the plane of the figure, *i.e.*, in the direction of the lengths of the filaments, so that the luminosity had its maximum brightness in line with the vertical plane through the axis of the collimator.

---

* G. Déjardin, 'Comptes Rendus,' vol. 172, p. 1483 (1921).

† F. Horton and A. C. Davies, 'Phil. Mag.,' vol. 41, p. 921 (1921).

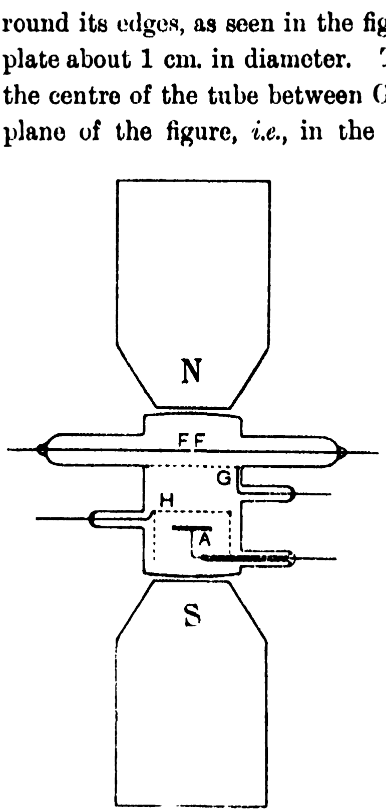


Fig. 1.

The second form of apparatus has already been described in detail by one of us in a paper on the excitation of the spectra of helium.[*] It was designed mainly for the investigation of low voltage arcs, and consists essentially of a spherical bulb fitted with three tungsten filaments, one of which was used at a time as the source of the bombarding electrons. The anode was a hollow platinum sphere 1 cm. in diameter. The distance of this from the glowing filament could be varied from 1 to 20 mm. The e.m.f. in the circuit was supplied by a 110 volts storage battery, and the potential difference across the discharge tube, and the current through it, could be regulated by means of series and parallel resistances. When a more delicate adjustment of the voltage drop across the tube was required, this was secured by means of a potentiometer arrangement which enabled the potential difference to be varied in steps of one-tenth of a volt.

In order to find the correction which has to be added on to the measured difference of potential between the anode and the glowing filament, so as to give the energy of the electrons in the stream, different arrangements of the electric fields were used in different cases. One of the methods employed with the second form of apparatus, described above, consisted in using one of the cold filaments and its leads as a third electrode, and ascertaining to what negative potential with respect to the negative end of the glowing filament this third electrode had to be charged in order to prevent electrons from the filament from reaching it when various potential differences were applied between the hot filament and the anode. This method was used to obtain corrections for ionisation current-voltage curves, and the delicate galvanometer which was employed as the detecting instrument was used at the same sensitiveness as when obtaining the corresponding ionisation current curves. Sometimes the determination of the correction was made using the platinum sphere as the testing third electrode and one of the cold filaments as the anode. It was found that, for any given conditions of pressure and filament temperature, the same value of the correction was obtained by each of the arrangements of electrodes indicated above. Moreover, the value of the correction found in this way was independent of the potential difference applied between the hot filament and the electrode which served as anode, from zero to the first critical voltage. The values of the corrections to be applied in total current curves were determined by finding the value of the negative potential difference which had to be applied between the negative end of the hot filament and the anode, in order to prevent electrons from reaching the anode in sufficient numbers to give an indication on the galvanometer when used at the same sensitiveness as in taking the corresponding total current-voltage curves. All the voltages marked on the curves, and referred to in the following pages, are the values of the electron energies obtained after applying the corrections to the measured potential differences.

---

[*] A. C. Davies, 'Roy. Soc. Proc.,' A, vol. 100, p. 599 (1922).

The argon used was carefully purified in the manner described in our earlier paper, and the same arrangements were employed for circulating the pure gas through the apparatus. The two forms of discharge tube were connected together, and the argon entered and left through U-tubes immersed in liquid air. Some of the experiments, more particularly those at the higher gas pressures, were carried out with the stopcock between the apparatus and the pump, pressure-gauge, etc., closed. Whenever this communication was opened the U-tube on the pump side of the apparatus was immersed in liquid air.

Before admitting argon to the discharge tubes, the residual gas was removed as completely as possible from the electrodes and glass walls by the method of baking and pumping with a mercury vapour pump, the filaments being maintained glowing during the process. This operation, as usual, occupied several days.

Single Ionisation and the Red Spectrum.

For moderate values of the intensity of the bombarding electron stream, it was found that, as the energy of the electrons was gradually raised, all the brighter lines of the red spectrum of argon appeared simultaneously, the point at which this occurred being a few volts higher than the ionisation voltage, as indicated by the rise in the ionisation or total current curves. In general, the first visibility of the lines was accompanied by a sudden large increase in the current through the gas, an increase due, no doubt, to the neutralisation of the space charge of the emitted electrons near the filament and the consequent large increase in the electron stream across the tube. After the neutralisation of the space charge had been effected, the voltage across the tube could be reduced considerably without the "red spectrum" lines of argon disappearing, and without there being a discontinuous fall in the current.

For small electron currents the limit to which the lines could be backed was only about 1 volt below the voltage at which they made their appearance, whereas, at the highest filament temperatures, and at moderate gas pressures, the difference between the voltages of appearance and disappearance of the "red spectrum" lines amounted to as much as 11 volts, the lower limit being far below the normal ionising voltage indicated in the curves. Over a considerable range of conditions, however, the limiting voltage for the disappearance of the "red spectrum" lines of argon had a definite fixed value, and a consideration of the current-voltage curves, taken simultaneously with the spectroscopic observations, serves to justify the interpretation of this definite limiting value for the disappearance of lines as the normal ionising voltage.

In general, the current-voltage curves showed the following characteristics: In curves for gradually increasing voltages there was a slight change of slope at about 11·5 volts, the value previously determined by us as the radiation voltage for argon. This was followed by a more marked change of slope at about 15 volts, a value which agrees well with our previous determination of the ionisation voltage. A few volts above this point, the actual voltage differing somewhat in different cases, there occurred a sudden sharp rise in the measured current, the increase being sometimes as much as a hundred times the original current. This rise indicates the increased electron emission already referred to, and was, with very few exceptions, coincident with the first appearance of spectrum lines. On gradually decreasing the applied potential difference a stage was eventually reached at which a correspondingly large decrease in current occurred, but the voltage at which this took place varied over as wide a range as the voltage of disappearance of the "red spectrum" lines. It was found, in fact, to be closely connected with the latter; for although the lines in certain instances ceased to be detectable visually a few tenths of a volt above the value at which the current fell, yet, on raising the potential difference slightly, lines were detected again provided the fall of the current had not occurred, whereas, if the fall had taken place, the lines could not be made to reappear until the voltage was increased to the value corresponding to the big rise of current in the increasing voltage curve. In view of the fact that, in a very large number of cases, the disappearance of the lines actually coincided with the large decrease of current, it is reasonable to suppose that the apparent discrepancy in the other cases is due to the lines becoming insufficiently intense for visual detection.

The value of the voltage, which was a definite limit for the backing of lines for a considerable range of conditions of pressure and filament temperature, agreed well with the value of the ionising voltage indicated in the curves for increasing voltages. Moreover, in the instances when the lines could be maintained below this limiting value, the decreasing voltage curves show a flattening which commences when this value is passed. These curves thus indicate that a change occurs in the sources of ionisation at this point. Such a flattening of the curve is to be anticipated at the normal ionisation voltage for argon, for below this voltage the production of ionisation by single impacts ceases, and ionisation can only occur by cumulative action. The fact that the spectrum lines of argon could be backed to values of the voltage below the usual limit of 15 volts, only for the larger pressures and for intense bombarding streams, supports the view that its maintenance below 15 volts is due to ionisation by cumulative action.

The continuous curves shown in figs. 2 and 3 illustrate the general characteristics of the current-voltage curves for *increasing* values of the voltage, while the curves drawn with a broken line illustrate the flattening of the curve for *decreasing* voltages when the spectrum lines are maintained by cumulative action (fig. 3), and the absence of such a bend when ionisation by cumulative action does not occur (fig. 2). The curves in fig. 2 represent the results of a series of observations of the variation in the total current between the hot filament and the anode, while the curves of fig. 3 represent the variation of the ionisation current alone. In order to obtain ionisation current curves with the tungsten filament apparatus the following method was used: Between the negative end of the hot filament and the leads of one of the other filaments (cold), a potential difference was applied which was adjusted to be just sufficiently great to prevent electrons from the hot filament from reaching the cold filament. The value of the potential difference necessary for this adjustment was found to be independent of the potential difference applied between the hot filament and the anode. A galvanometer was included in the circuit between the filaments, and observations of the current were made as the voltage between the hot filament and anode was increased. When ionisation and total current measurements were taken simultaneously, as was sometimes the case, there was no essential difference in the characteristics of the curves obtained. In the two sets of observations represented above the pressure of argon was the same, namely, 0·784 mm., but a much more intense electron stream was employed in the case of the curves of fig. 3 than in those of fig. 2.

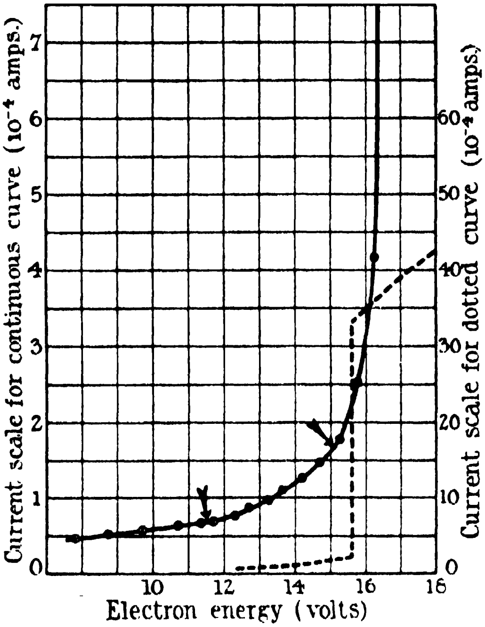


Fig. 2.

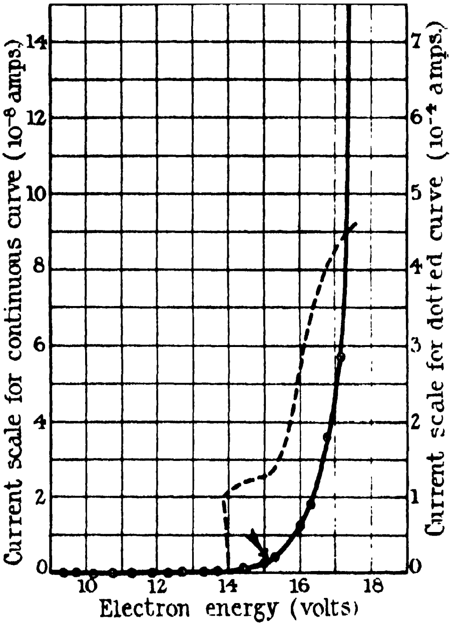


Fig. 3.

The extent to which ionisation was detected below the voltage which corresponds to the normal ionisation of argon, in the curves for gradually increasing values of the voltage, varied very much in different experiments. That the effect detected was really ionisation, and not simply a photoelectric effect of radiation on the negatively charged electrode, was suggested by the fact that the bend is as apparent in the total current curves as in the ionisation curves. It is inconceivable that a current due to a photoelectric emission of electrons, resulting from the radiation produced by the collisions of the primary electrons with gas atoms, could be comparable with the current due to the primary electrons themselves, whereas such an occurrence is readily accounted for if positive ions are produced, since the presence of these gives rise to an increase in the current due to the primary emission itself. That the view is justified was shown in two ways: Firstly, by using the platinum sphere instead of the cold filament as the electrode from which photoelectric emission could occur. It was found that the magnitude of the effect was unchanged by this alteration. Secondly, in a circuit between the hot filament and one of the cold filaments, the potential difference was so adjusted as to allow a small electron current from the hot filament to pass to the cold filament when a potential difference, which could be varied, was applied between the hot filament and the platinum sphere. It was found that an increase in the electron current to the cold filament occurred simultaneously with the bend in the total current-voltage curve, thus indicating that an increase in the emission from the hot filament had occurred, which could only be brought about by the presence of positive ions.

The cause of the first bend in total current curves and in ionisation current curves must thus be either ionisation of impurity or ionisation of the argon by cumulative action. That the latter is, at any rate, not the main factor is well illustrated by the curves already given in fig. 2, for the curve for increasing voltages shows the 11·5 volts bend quite clearly, whereas the curve for decreasing voltages indicates, according to the reasoning given earlier, that ionisation of argon by cumulative action was not occurring to any extent in this case, even with the increased intensity of the bombarding electron stream which results from neutralisation of the space charge. Hence it appears that the ionisation detected below the normal ionising value in curves for increasing voltages is mainly ionisation of impurity.

Other evidence which points to this conclusion is that the bend was found to be more marked when observations were taken in argon which had been in the apparatus for some time, and in which the filament had been hot for a long period, than when observations were made with argon streaming through the apparatus after submitting the latter to a prolonged pumping. This indicates, moreover, that the impurity arises in the apparatus itself, and is not admitted with the argon. The amount of impurity present must always have been very small, on account of the long continued outgassing to which the apparatus had been subjected, and to the fact that the argon in the apparatus was frequently changed, even when it was not continuously streaming through during the experiments. The only impurity lines which were seen in the spectrum of the luminosity were the H_α and H_β lines, which were visible on some occasions, and a few lines whose wave-lengths agreed with those of certain of the brightest lines of the tungsten arc spectrum. These latter lines were only observed with very brilliant discharges when currents of about 0·5 ampere were passing through the tube. It must be borne in mind, however, that the spectroscopic test for impurities in argon is not one of extreme delicacy. On some occasions, especially with the apparatus of fig. 1, it was thought that traces of bands could be seen in the spectrum when the conditions were such that the argon lines were only faintly visible. It seemed possible that these might be the carbon oxide spectrum, but the appearance was so extremely faint that we could not be certain from visual observations, and we were unable to obtain a photographic record of the presence of bands even with 8 hours' exposure.

With a very small amount of impurity present in the argon, it would not be expected that sufficient ionisation of it, by direct electron impact, to give an indication in the measured currents would take place; but if this impurity is ionisable by the argon radiation, an indication of ionisation might be expected when this radiation is produced.* Thus the value of the voltage at which ionisation is detected in the increasing voltage curves should be the minimum value for the production of resonance radiation from argon. The value of the resonance voltage deduced in this manner (11·5 volts, fig. 2) agreed with that found in our previous determinations.

When the pressure of the argon was of the order of 1 mm., and when the electron bombardment was intense, ionisation of the gas by cumulative effects took place, probably as the result of electron impacts on argon atoms which were in an abnormal state owing to the absorption of resonance radiation from neighbouring atoms. With the filament as hot as it was safe to make it, the voltage across the discharge tube could be reduced below the minimum radiation voltage for argon without the "red spectrum" lines disappearing. This effect is similar to that obtained by one of the writers with helium,† and is presumably to be explained by the frequency of electron collisions with argon atoms being so great that, on recombination occurring, the returning electron is unable to fall right back to the normal orbit before being again ejected by another impact. It would be expected that the lowest voltage at which all the lines of the red spectrum (excluding those in the extreme ultra-violet) could be maintained would be the minimum voltage required to ionise an atom which has absorbed resonance radiation, *i.e.*, 15·1 volts − 11·5 volts = 3·6 volts. The lowest value to which the voltage could be backed in the present experiments was 6·5 volts, but it must be borne in mind that there are two factors which tend to make the ionisation diminish very rapidly as the minimum value of the electron energy which can produce it—3·6 volts—is approached, viz.:—(*a*) the decreasing efficiency of the electrons in producing ionisation as their energy approaches the critical value; (*b*) the increased proportion of recombination which occurs between positive ions and electrons as the intensity of the electric field is reduced. Owing to both these causes, the number of positive ions reaching the neighbourhood of the filament diminishes until an effective neutralisation of the space charge is no longer maintained. When this stage is reached the electron current across the tube decreases abruptly, and the electron bombardment is then not sufficiently intense to prevent the reversion of the abnormal atoms to the normal condition, so that ionisation ceases.

---

* F. Horton and D. Bailey, 'Phil. Mag.,' vol. 40, p. 440 (1920).

† *Loc. cit.*

A striking series of changes in the luminosity of the "red spectrum" lines was observed when the voltage across the tube was gradually reduced from about 25 volts. To begin with, the brightness gradually diminished, but following this there was a noticeable *increase* in the intensity of most of the bright lines as the voltage fell from about 19 volts to about 16 volts, after which the lines diminished in intensity, and finally disappeared. The bright lines λ 7067·5 and λ 6965·8 in the red were not among those whose intensity varied in this way, for they seemed to decrease continuously in brightness as the voltage was lowered. The enhancement with decreasing voltage between 19 and 16 volts was particularly striking in the cases of the lines λ 4510·9, λ 4345·3, λ 4272·3, and lines farther in the blue and violet.

Another point in connection with the behaviour of different lines in the red spectrum under different conditions is worth recording. For small intensities of electron bombardment, the lines in the red and orange parts of the spectrum were only faintly visible even when the lines in the blue were fairly bright. As the current heating the filament was increased, the lines in the red, orange, yellow and green regions increased in brightness far more than the blue and violet lines, and many more lines became visible in these regions. This brightening was particularly striking in the case of the lines λ 7067·5 and λ 6965·8, which were the brightest lines in the red for intense electron emissions, but which for smaller values of the electron emission were hardly visible even when other red lines could be seen clearly.

In general the voltage at which the luminosity first appeared was between 17 volts and 18 volts, but it was found that when a magnetic field was used with the apparatus of fig. 1 the minimum voltage at which a bright luminosity* was produced was 20 volts or more. This delaying of the production of a bright glow in the gas until a higher potential difference had been established is probably due to the concentration of the electron stream into a narrow beam, which increased the chance of the recombination of positive ions and electrons when ionisation of the gas took place, and so delayed the stage at which positive ions reached the neighbourhood of the filament in sufficient numbers to produce an effective neutralisation of the space charge.

---

* In some cases with the apparatus shown in fig. 1 a very faint glow was seen at a lower voltage than that at which the sudden large increase of current occurred, and at which the bright glow appeared. This faint glow was presumably due to the recombination which took place before neutralisation of the space charge was effected.

Although about 17 volts was, in general, the lowest voltage at which luminosity was seen when the energy of the electron stream was raised gradually, it was found that if the filament were maintained very hot and *suddenly* the potential difference was switched on, the luminosity would appear (with pressures greater than about 1 mm.) at considerably lower voltages. The lowest voltage at which this effect was obtained was 12·5 volts, the reading of the voltmeter in the presence of the glow being several volts lower than this. When the arc strikes in this way, at a voltage only slightly higher than the minimum radiation voltage, the ionisation of the gas is produced by cumulative effects and positive ions are formed from the start in sufficient numbers to prevent the limitation of the electron current by the mutual repulsion of the emitted electrons, whereas when the voltage is increased gradually this limitation of the electron current is effected before ionisation is produced.

Multiple Ionisation and the Blue Spectrum.

In general the lines of the blue spectrum of argon were not present in the spectrum of the luminosity when it first appeared, but required for their stimulation a higher potential difference across the tube. This is to be expected on the view that these lines result from the recombination of electrons with argon atoms from which two or more electrons have been removed. As the voltage was gradually raised these lines did not appear suddenly, as was usually the case with the "red spectrum" lines, but they were very faint at first and brightened fairly rapidly as the voltage was further increased. The lowest voltage at which they were seen, for increasing values of the voltage, depended to a certain extent upon the intensity of the electron stream, and increased as that intensity was diminished.

Two limiting values of the voltage at which any lines of the blue spectrum could be detected by visual observations were obtained under different conditions. The greater of these, which if Stark's view is correct must be the minimum voltage necessary for the simultaneous removal of two electrons from an argon atom, was about 34 volts, while the lower one, which it would follow is the minimum voltage necessary for the removal of a second electron from an already ionised argon atom, was about 19 volts. It was difficult to fix the upper limit with certainty by visual observations because of the gradual manner of the appearance and disappearance of the "blue spectrum" lines. Moreover, it was impracticable to increase the visibility of the lines at voltages slightly above the limit, by raising the temperature of the filament because, except at the lowest pressures (*e.g.*, 0·001 mm.) this usually resulted in the production of multiple ionisation by cumulative effects. The limiting values of the voltage for the production of lines of the blue spectrum were therefore investigated photographically, giving long exposures so as to make the detection of lines of feeble intensity more probable. The current-voltage relations were also investigated with a view to determining the voltage at which double ionisation of the argon atom occurred, so as to test the view, already referred to, of the connection between the blue spectrum and multiple ionisation of the atom.

The photographic method of investigation consisted in taking series of photographs of the spectrum of the luminosity at different voltages at intervals of 1 or 2 volts, keeping the total current through the tube constant at all the voltages throughout a series, and examining the plates for the presence of any lines belonging to the blue spectrum only. Among the lines of the blue spectrum which were the first to become visible as the voltage was increased were λ 4430·4 and λ 4426·2, and these lines and others are distinctly visible on all the plates taken with fairly intense electron streams for voltages higher than 19. On the plate taken at 19 volts these two lines can just be detected but they are extremely faint, and at 18·5 volts they are not present at all. A comparison of the plates at different voltages shows that the intensity of the "blue spectrum" lines increased as the voltage was raised from 19 volts to about 24 volts, after which it did not increase further until the voltage reached 34. From 34 volts up to 42 volts the "blue spectrum" lines increased very much in intensity and became more prominent than the lines of the red spectrum. The increase of intensity was first noticeable on the plate at 34 volts. In order to keep the current across the tube the same at all the voltages in a series, it was necessary to reduce the intensity of the electron stream as the voltage was raised. The change in the intensity of the "blue spectrum" lines after 34 volts must, therefore, have been due to the increase in the electron energy and not to an increased intensity of bombardment. Hence the results of the photographic investigation confirm the results of the visual spectroscopic observations and show that there are two limiting voltages associated with the production of "blue spectrum" lines, namely, 19 volts and 34 volts.

The curves given in figs. 4 and 5 illustrate the results obtained in the investigation of current-voltage relations at voltages greater than the normal ionising voltage. The results represented by fig. 4 were taken in argon at a pressure of 0·75 mm., and with the anode 7 mm. from the glowing filament. The curve shows a distinct break at 30 volts. Fig. 5, on the other hand, expresses the results obtained at a pressure of 0·024 mm., and with the anode and filament only 2 mm. apart. This curve is rather different in form from that of fig. 4, for the current continues to increase considerably with increasing voltage for a much bigger range than in the latter case. There is no indication of a break at 30 volts, but a distinct break occurs at about 34 volts. This difference in the curves is accounted for by the different value of the mean free path of the electrons in the two cases. In the circumstances of fig. 4 each electron would make several collisions between the filament and the anode, so that a bend would be expected to occur in the curve at twice the normal ionising voltage. In the case of fig. 5, the mean free path of the electrons was larger than the distance between the electrodes, so that very few of the electrons were likely to collide twice before reaching the anode. A bend in the curve at twice the normal ionising voltage would therefore not be expected in this curve. Hence it is reasonable to suppose that the bend at 30 volts in the curve of fig. 4 is due to a considerable proportion of the electrons in the stream twice producing single ionisation, while the bend in the curve at 34 volts in fig. 5 is due to the production of doubly ionised argon atoms by the simultaneous removal of two electrons from the normal atom.

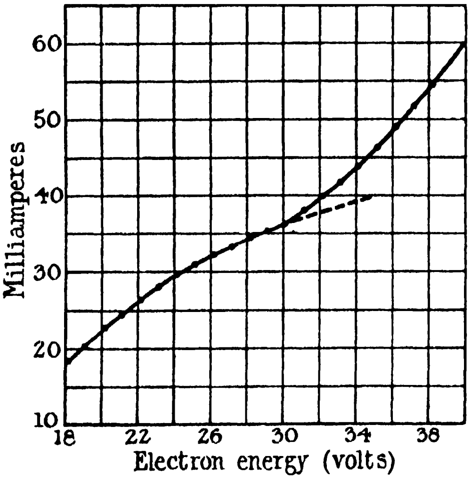


Fig. 4.

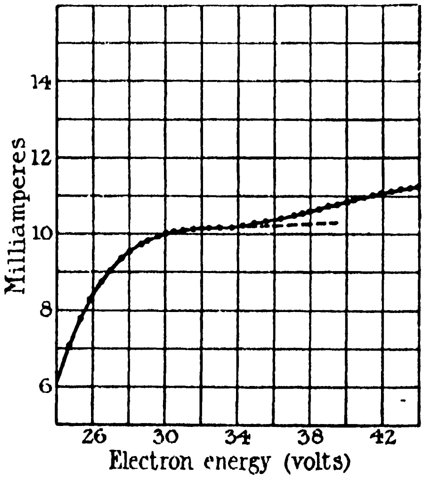


Fig. 5.

The position of the bend in the curve of fig. 4 thus confirms the value of the normal ionising voltage determined by other methods, while the fact that the value of the double ionisation voltage obtained from curves such as that of fig. 5 agrees, within the limits of experimental error, with the greater of the two limits for the appearance of lines of the blue spectrum supports the view that some of these lines result from the recombination of doubly ionised argon atoms and electrons. The difference between the double ionisation voltage and the normal ionising voltage must be equal to the voltage which is necessary for the removal of a second electron from an already ionised atom, *i.e.* (34 volts − 15 volts =) 19 volts, must be the minimum voltage at which two electrons can be removed from an argon atom by two electron collisions. This value agrees with the lower of the two limits obtained in connection with the appearance of "blue spectrum" lines, and hence supports the view that some of the lines of this spectrum are associated with the presence of argon atoms carrying two positive charges.

When exceedingly intense bombarding electron streams were employed the lines of the blue spectrum, like those of the red spectrum, could be detected at voltages considerably below the usual limits associated with their production. For example, in the instances when the "red spectrum" lines appeared as the voltage was raised to 13 volts or 14 volts, the "blue spectrum" lines were plainly detected in the first flash of the luminosity, but since the potential difference across the tube always fell several volts when the glow appeared, the "blue spectrum" lines did not generally remain; they were visible for an instant only, and were then not seen until the voltage was raised. The lowest voltage at which they were detected on raising the potential difference was 11 volts.

As a result of his experiments on canal rays in argon, Stark has concluded that the blue spectrum of this gas includes radiations which are associated with the presence of trebly charged argon atoms, as well as radiations associated with doubly charged atoms. This suggests that different lines of the blue spectrum require different voltages for their production by electron bombardment. The series of photographs were therefore examined carefully, with a view to ascertaining which lines were associated with atoms having more than two charges. The positions of most of the brighter lines of the blue spectrum, between λ 4888·9 and λ 4014·0, given in the Tables in Kayser's "Handbuch der Spectroscopie," were ascertained from one of the photographs taken at about 46 volts, and the intensity of each line on a series of plates was then investigated. Tables I and II, given below, illustrate the results obtained in the cases of the lines which Stark has classified as "bivalent" and "trivalent" respectively, while Table III gives the results of the investigation of several "blue spectrum" lines between λ 4888·9 and λ 4014·0 which are not classified by Stark.

In each Table a rough estimate of the intensity of the lines at different voltages is given. The presence of any of these lines below 34 volts (the limit at which the simultaneous removal of two electrons by a single collision can occur) is to be attributed to the occurrence of multiple ionisation by cumulative effects. The increased brightness of many of the lines when the voltage was raised beyond 34 volts is illustrated by the numbers given.

It will be seen from the Tables that there are considerable differences in the minimum voltages at which different " blue spectrum " lines were detected, and that so far as any classification of the lines can be made on the basis of the minimum voltage necessary for their production, this classification does not agree with that given by Stark. It would appear that of the lines classified by Stark, those requiring most energy for their production are λλ 4202·1, 4082·6, 4076·9, 4043·0 and 4014·0, and that these are the " trivalent lines " or, alternatively, that all the lines requiring 38 volts or more for their production are " trivalent," in which case this class includes most of the lines classified by Stark as " bivalent." It must, however, be remarked that if 38 volts is sufficient to produce trebly charged positive ions, these must result from collisions with already ionised argon atoms, for it is exceedingly unlikely that the energy necessary to remove three electrons from a normal argon atom is only 4 volts in excess of that required to remove two

Table I.—Stark's " Bivalent " Lines.

Wave-length.	Intensities.—Electron Energy (volts).										
	42.	40.	38.	36.	34.	32.	30.	28.	26.	24.	20.
4309·3	3	1	0·5	—	—	—	—	0·5	0·5	0·5	—
4300·8	1	0·5	0·5	—	—	—	—	—	—	—	—
4283·0	3	1	0·5	—	—	—	—	—	—	—	—
4277·7	8	6	5	2	2	2	2	2	2	2	—
4237·3	3	1	0·5	—	—	—	—	—	—	—	—
4228·7	8	5	4	1	0·5	0·5	0·5	0·5	0·5	0·5	0·5
4202·1	—	—	—	—	—	—	—	—	—	—	—
4132·0	3	3	2	—	—	—	—	—	1	2	—
4082·6	—	—	—	—	—	—	—	—	—	—	—
4072·2	3	2	1	—	—	—	—	—	—	—	—
4043·0	—	—	—	—	—	—	—	—	—	—	—
4014·0	—	—	—	—	—	—	—	—	—	—	—

Table II.—Stark's " Trivalent " Lines.

Wave-length.	Intensities.—Electron Energy (volts).							
	42.	40.	38.	36.	34.	32.	30.	28.
4434·1	3	1	0·5	0·5	0·5	0·5	—	0·5
4222·8	2	1	—	—	—	—	—	—
4218·8	2	1	—	—	—	—	—	—
4156·3	1	0·5	—	—	—	—	—	—
4104·1	4	3	2	1	1	0·5	0·5	0·5
4076·9	—	—	—	—	—	—	—	—

Table III.—Other "Blue Spectrum" Lines.

Wave-length.	Intensities.—Electron Energy (volts).												
	42.	40.	38.	36.	34.	32.	30.	28.	26.	24.	20.	19.	18·5.
4888·9	2	1	1	1	1	—	—	—	—	—	—	—	—
4866·1	0·5	—	0·5	—	—	—	—	—	—	—	—	—	—
4847·9	3	2	2	1	0·5	—	—	—	—	—	—	—	—
4806·2	6	3	3	3	2	1	0·5	1	0·5	0·5	0·5	—	—
4765·0	6	3	4	3	1	0·5	0·5	0·5	0·5	0·5	—	—	—
4736·0	5	3	3	2	1	0·5	0·5	0·5	0·5	0·5	—	—	—
4727·0	5	3	4	2	1	0·5	0·5	0·5	0·5	0·5	—	—	—
4658·0	7	4	4	3	1	0·5	0·5	—	—	—	—	—	—
4637·4	2	1	0·5	—	—	—	—	—	—	—	—	—	—
4609·7	7	4	4	3	1	1	1	2	1	1	—	—	—
4590·1	6	4	3	2	1	0·5	0·5	0·5	0·5	0·5	—	—	—
4579·5	6	4	4	3	1	0·5	0·5	0·5	0·5	0·5	0·5	—	—
4545·3	6	5	5	4	2	0·5	0·5	1	0·5	1	0·5	—	—
4503·2	1	0·5	—	—	—	—	—	—	—	—	—	—	—
4498·7	1	—	—	—	—	—	—	—	—	—	—	—	—
4491·2	1	1	0·5	—	—	—	—	—	—	—	—	—	—
4488·4	—	—	—	—	—	—	—	—	—	—	—	—	—
4482·0	8	7	5	3	1	1	1	2	0·5	1	—	—	—
4431·2	3	3	3	2	1	—	—	—	—	—	—	—	—
4430·4	8	7	6	4	3	1	1	2	1	2	1	0·5	—
4426·2	10	10	10	8	5	2	2	3	2	3	2	0·5	—
4401·2	8	7	7	6	4	2	2	2	2	2	1	0·5	—
4400·3	5	4	3	3	1	0·5	—	—	—	—	—	—	—
4379·8	7	7	7	5	3	1	2	2	2	2	1	—	—
4371·5	—	—	—	—	—	—	—	—	—	—	—	—	—
4370·9	8	8	7	5	3	1	2	2	2	2	1	—	—
4362·2	3	2	1	—	—	—	—	—	—	—	—	—	—
4352·4	5	4	4	3	1	—	—	0·5	—	—	—	—	—
4348·1	10	10	10	10	8	5	4	5	3	4	3	0·5	0·5
4332·2	1	1	1	—	—	—	—	—	—	—	—	—	—
4331·3	1	0·5	0·5	—	—	—	—	—	—	—	—	—	—

electrons simultaneously from such an atom, namely, 34 volts. The significance to be attached to 38 volts as a critical electron energy—if this voltage is a genuine limit for the production of trebly ionised atoms by cumulative effects—is doubtful, because of the various possible ways in which the energy necessary for the removal of three electrons might be acquired from the combined effects of two or three electron collisions.

Conclusion.

If the square roots of the voltages connected by the quantum relation with the frequencies of the various K X-ray absorption limits are plotted against the corresponding atomic numbers, it is found that the square root of the accepted value of the ionisation voltage of helium falls on the curve at the point corresponding to atomic number 2. This fact has been interpreted as indicating that the ionisation voltage of helium corresponds to the K X-ray absorption limit of this element. A similar relation has been shown by

Mohler and Foote[*] to hold in the case of the ionisation voltage of neon (16·7 volts)[†] and the L absorption limits. It might, therefore, be anticipated that in the case of argon a similar connection would hold for its ionisation voltage and the voltages corresponding to one of the M absorption limits of other elements. Except in the case of the elements of high atomic number, however, very little information with regard to M absorption limits is available. In the cases of the elements molybdenum (atomic number 42), copper (atomic number 29), and potassium (atomic number 19) determinations of the voltages corresponding to the M radiations have recently been made.

Richardson and Bazzoni[‡] found that, in the case of molybdenum, soft X-rays were produced when the applied voltage was about 356 volts, and they classified the radiation they obtained as belonging to the M series. In fig. 6, the square roots of the voltages connected by the quantum relation with the frequencies of the M_α, M_β and M_γ lines are plotted against the corresponding atomic numbers for the elements for which M X-ray data are available.[§] It will be seen that the prolongation of the M_γ line passes through the molybdenum point determined by Richardson and Bazzoni. It seems probable that the voltage required to excite M radiations by electron bombardment is that corresponding to the frequency of one of the M absorption limits. Hence it is reasonable to assume that one of the M absorption limits for molybdenum is not far from 356 volts. The curve drawn with a broken line in fig. 6 joins the points obtained by plotting the square root of the voltage corresponding to the frequency of the M absorption limit (ν_{M_a}) against the atomic number (A) for the cases of argon and molybdenum, taking these voltages to be 15·1 and 356 respectively. If the relation $\sqrt{(\nu_{M_a})} = k\mathrm{A}+c$ holds between molybdenum and argon, then the value of the voltage corresponding to ν_{M_a} for any element intermediate to these should be obtainable from this line. The value deduced for potassium from this curve is 20·3 volts, which agrees within the limits of experimental error with one of the critical points recently found by Mohler and Foote[||] in the radiation curve for this element. Two breaks were found by these experimenters in the radiation curve, one at 20 ± 1 volts and another more marked break at 23 ± 1 volts. A break at the latter point was also obtained in their ionisation curves. Within the limits of experimental error, the difference between these two voltages agrees with the normal ionising voltage of potassium found by these experimenters, viz., 4·3 volts. This fact suggests that the lower of the two voltages given above is the one which really corresponds to the M absorption limit for potassium, and that the greater of the two voltages given above is that required to remove two electrons simultaneously from the potassium atom, the outer more loosely attached electron, and one from the M ring of electrons. On this view, the occurrence of a break in the ionisation curves at 23 ± 1 volts is readily explained, and the agreement between the lower value given above and that determined from the curve of fig. 6 is very satisfactory.

---

[*] F. L. Mohler and P. D. Foote, 'Journal of the Opt. Soc. of America,' vol. 5, p. 328 [1921].

[†] F. Horton and A. C. Davies, 'Roy. Soc. Proc.,' A, vol. 98, p. 124 (1920).

[‡] O. W. Richardson and C. B. Bazzoni, 'Phil. Mag.,' vol. 42, p. 1015 (1921).

[§] W. Duane, 'Bulletin of Nat. Research Council,' vol. 1, p. 383 (1920).

[||] F. L. Mohler and P. D. Foote, 'Phys. Rev.,' vol. 18, p. 94 (1921).

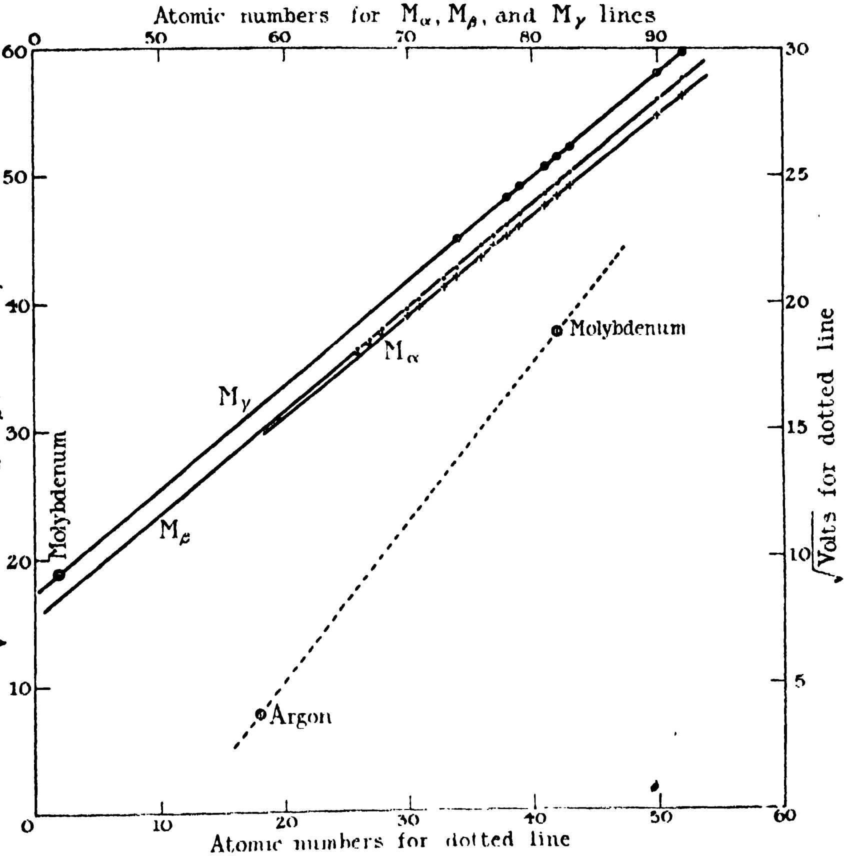


Fig. 6.

In the case of copper, the other element between molybdenum and argon for which M series X-ray data are available, the value of the voltage corresponding to the frequency of the M absorption limit deduced from the curve is about 116 volts. This value agrees within the limits of experimental error with one of the critical points found by Kurth,* but ascribed by him to a hypothetical N series. Kurth obtained another break in his radiation curve for copper, at about 300 volts, and this he attributes to the M series of copper. This value is, of course, very far from falling on the curve connecting Richardson's molybdenum value and our own argon value. Kurth admits, however, that the value attributed by him to an N series may really correspond to an M series.

---

* E. H. Kurth, 'Phys. Rev.,' vol. 18, p. 461 (1921).

In the case of iron also, Kurth obtained breaks in the radiation curve, which he attributed to N and M series respectively. Neither of these values, however, fall very near the molybdenum-argon line in fig. 6. Experiments to investigate the voltages necessary for the excitation of the M radiations of other elements in this region are at present in progress in this laboratory.

Summary.

The ionisation of argon by electron collisions has been investigated by spectroscopic observations taken in conjunction with current-voltage curves. The authors' previous determinations of the minimum electron energies for the production of resonance radiation (11·5 volts) and ionisation (15·1 volts) have been confirmed and, in addition, the energy necessary to remove two electrons simultaneously from a normal argon atom has been found to be 34 volts.

It has been found that many lines in the blue spectrum of argon appear simultaneously with the occurrence of double ionisation, but that others need a greater electron energy for their stimulation, and therefore probably require the removal of more than two electrons from the normal atom for their production.

Certain of the lines of the blue spectrum have been classified by Stark as "bivalent" and certain others as "trivalent," but this classification is not confirmed by the results of the determination of the minimum electron energies required for their stimulation.

So far as can be judged from the limited data available, it appears probable that the ionisation voltage (15·1 volts) is related to one of the M X-ray absorption limits of the various elements in a manner analogous to that in which the helium ionisation voltage is related to the K absorption limits and the neon ionisation voltage (16·7 volts) to the L absorption limits. Thus it seems reasonable to suppose that the frequency with which the ionisation voltage of argon is connected by the quantum relation is that of one of the M absorption limits for argon.

The authors wish to express their thanks to the Radio Research Board of the Department of Scientific and Industrial Research for the means of purchasing most of the apparatus used in this investigation.

The Molecular Scattering of Light in Vapours and in Liquids and its Relation to the Opalescence observed in the Critical State.

By K. R. RAMANATHAN, M.A., Demonstrator in Physics, H.H. the Maharajah's College, Trivandrum; University of Madras Research Scholar, Calcutta.

(Communicated by Dr. Gilbert T. Walker, F.R.S. Received June 19, 1922.)

1. *Introduction.*

It has long been known that in the immediate vicinity of the critical state, many substances exhibit a strong and characteristic opalescence. In recent years, the phenomenon has been studied by Travers and Usher[*] in the case of carefully purified CS_2, SO_2, and ether, by S. Young,[†] by F. B. Young[‡] in the case of ether, and in a quantitative manner by Kammerlingh Onnes and Keesom[§] in the case of ethylene. An explanation of the phenomenon on thermodynamic principles as due to the accidental deviations of density arising in the substance was put forward by Smoluchowski.[||] He obtained an expression for the mean fluctuation of density in terms of the compressibility of the substance, and later, Einstein[¶] applied Maxwell's equations of the electromagnetic field to obtain an expression for the intensity of the light scattered in consequence of such deviations of density. He showed that the fraction α of the incident energy scattered in the substance per unit volume is

$$\frac{8\pi^3}{27} RT\beta \frac{(\mu^2-1)^2(\mu^2+2)^2}{N\lambda^4}. \tag{1}$$

In this, R and N are the gas constant and Avogadro's number per gram-molecule, β is the isothermal compressibility of the substance, μ is the refractive index and λ is the wave-length of the incident light. Keesom[**] tested this formula over a range of 2·35° above the critical point of ethylene and found good agreement except very close to the critical point.

Of an altogether different order of intensities is the phenomenon of light-scattering in gases and vapours under ordinary pressures observed experimentally by Cabannes[††] and the present Lord Rayleigh.[‡‡] The effect studied by these investigators is, in comparison with the phenomenon of critical opalescence, feeble, and requires powerful illumination and special arrangements to enable it to be observed in the laboratory. The measurements made by them show that the magnitude of the effect[*] is in substantial agreement with the theory of molecular scattering of light developed by the late Lord Rayleigh, according to which the extinction-coefficient α is given by

---

[*] 'Roy. Soc. Proc.,' A, vol. 78, p. 247 (1908).

[†] *Ibid.*, p. 262.

[‡] 'Phil. Mag.,' vol. 20, p. 793 (1910).

[§] 'Roy. Soc. Proc. Amsterdam,' 1908.

[||] 'Ann. der Physik,' vol. 25, p. 205 (1908).

[¶] 'Ann. der Physik,' vol. 33, p. 1275 (1910).

[**] 'Ann. der Physik,' vol. 35, p. 591 (1911).

[††] 'Comptes Rendus,' vol. 160, p. 62 (1915); and 'Ann. de Physique,' vol. 15 (1920).

[‡‡] 'Roy. Soc. Proc.,' A, vol. 94, p. 453 (1918).

$$\frac{32\pi^3}{3}\frac{(\mu-1)^2}{n\lambda^4}, \qquad (2)$$

where n is the number of molecules per unit volume. The scattering of light in passage through *dust-free liquids* has also been observed by the present Lord Rayleigh[†] and W. H. Martin,[‡] and it was found by the former that the intensity of scattering in the liquid studied (ether) was considerably smaller than that given by the formula (2), though, owing to the greater density of the liquid, the absolute magnitude of the effect is many times larger than in the vapour.

It is important to notice that these three phenomena, namely, critical opalescence, scattering of light by gases and vapours, and scattering by liquids, which are of such different orders of magnitude, and apparently so distinct, stand in reality in the closest relationship to each other. It has been pointed out by Prof. C. V. Raman[§] that the observed scattering power of liquids under ordinary conditions agrees with that calculated from the Einstein-Smoluchowski formula; and further, that this law which reduces to Rayleigh's formula in the special case of a gas obeying Boyle's law, is much more generally applicable than the latter and should be regarded as the fundamental relation determining the magnitude of light-scattering in all fluid media. It thus becomes a matter of great importance to make a thorough test of the Einstein-Smoluchowski formula over the widest possible range of experimental conditions. A series of investigations with this object has been undertaken at Calcutta.

The present paper describes the results of a study made of the scattering of light in ether, both in the liquid and in the saturated vapour phases at different temperatures from 30° up to the critical point 193·6°, and also in the gaseous phase above the critical temperature up to 217° C. Ether was chosen as the first substance to be investigated, because it possesses a high refractivity and hence a large scattering power, is easily obtained pure, and can withstand the action of heat and moderately intense light without appreciable change. Moreover, the classical work of Ramsay and Young* on the pressure-volume relations of the substance over a wide range of temperature provides the necessary data for calculation.

---

[*] 'Roy. Soc. Proc.,' A, vol. 95, p. 155 (1918).

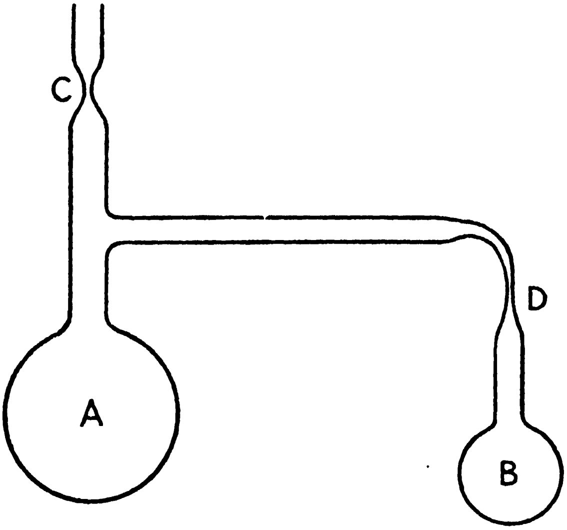
[†] 'Roy. Soc. Proc.,' A, vol. 95, p. 95 (1918).

[‡] 'Journal of Physical Chemistry,' 1919.

[§] 'Nature,' November 10, 1921, and 'Roy. Soc. Proc.,' A, April, 1922. See also "Molecular Diffraction of Light," by C. V. Raman, published by the Calcutta University Press, 1922, pp. 12 and 60.

2. *Experimental Arrangements.*

The ether used was Merck's pure ether, redistilled four or five times *in vacuo.* It was contained in a thick, uniform glass bulb of about 3 cm. external diameter. The necessary quantity of ether was introduced into it in the following manner. A glass apparatus of the form shown in the figure was made. The bulbs and connecting tube were well heated over a Bunsen flame and the air exhausted by means of a pump. Nearly four times the required quantity of ether was introduced into the bulbs, and in order to remove the dissolved air, the apparatus was exhausted until more than half the ether had evaporated and it was then sealed off from the pump at C. All the ether was transferred to the bulb A, and a portion was distilled over into B by placing A in a vessel of warm water at 45–55° C. and B in a vessel of water at the room-temperature (about 30° C.). When B was nearly half filled, its inside was washed out with the distilled ether, which was then transferred back to A without allowing the undistilled ether to come into B. This process was repeated four or five times in order to get the ether perfectly dust-free, and finally, when the required quantity had collected in B, the bulb was sealed off at D.



---

* 'Phil. Trans.,' A, vol. 178, p. 57 (1887).

The outside of the bulb was painted dead black, leaving three windows, two at the sides to let in and let out light, and one at the base for the observation of the scattered light. It was then mounted in the centre of a thick iron cross-tube, blackened inside, similar to the one employed by Lord Rayleigh (*loc. cit.*). The cross-tube was nearly 1·5 inches external diameter, and each arm of the cross was 7 inches long. Its ends were closed with glass windows, and it could be heated to any temperature up to 250° C. by passing an electric current through a coil of Eureka wire wound uniformly over it. The temperature could be maintained steady within less than half a degree for hours. The distribution of temperature was very uniform in the middle of the tube. The tube was mounted horizontally and sunlight was concentrated at the centre of the bulb by means of a lens. Different apertures could be introduced in front of the lens to regulate the quantity of light. Suitable diaphragms were also introduced inside the cross-tube to prevent diffuse light from reaching the eye. With this arrangement the track of the light in the ether vapour was easily visible through the window at the base of the bulb at room temperature when an aperture of 1 inch diameter was in front of the lens.

The intensity of the scattered light was measured by comparing it with the intensity of a similar track in distilled ether kept outside in a rectangular glass bottle. A lens of the same focal length as the one mentioned before was used to produce the track, and suitable apertures were introduced here also to regulate the brightness. The image of the second track was brought into the same field of view by means of two totally reflecting prisms. The prisms exercised very little absorption and the colour-match of the two tracks was practically perfect. The photometry was done with an Abney rotating sectored disc. In order to minimise the heating effect of sunlight, the apertures were made of the smallest size consistent with the visibility of the track, and the light was kept on for as short a time as possible. Temperatures were maintained steady for nearly half an hour before readings were taken.

A correction has to be made for the loss of light on reflection at the glass windows and at the walls of the bulb. It will be noticed that there are six glass-air surfaces and two glass-vapour or glass-liquid surfaces to be taken into account in the case of the bulb, and six glass-air surfaces and two glass-liquid surfaces in the case of the ether bottle used as standard. There would thus be an approximate compensation. To test if it is so, and to allow for any absorption which the right-angled prisms might exercise, a comparison was made of the light scattered by the ether in the bulb at room temperature and that scattered by the standard. The two were found to be equal within 2 per cent.

3. *Results.*

The observed data for the intensity of the scattering by the saturated vapour in terms of liquid ether as a standard are shown in Table I. The theoretical results indicated by the Einstein-Smoluchowski formula are shown in the third column after applying the correction for imperfect polarisation of the scattered light, as explained below in Art. 5. In order to calculate the theoretical value, it is necessary to know the compressibility, β or $-(1/v\ dv/dp)_t$, and the refractive index, μ, of the saturated vapour at the respective temperatures of observation, and similar data for the liquid used as a standard at 35° C. The compressibilities, $-(1/v\ dv/dp)_t$, of the vapour at different temperatures were determined by plotting Ramsay and Young's pressure-volume data on graph-paper and drawing tangents at the points of saturation. The values of μ were obtained by the application of Lorentz's refraction formula $(\mu^2-1/\mu^2+2) = k\rho$, where k has the value 1/3·317 for ether.

It will be seen that throughout there is a good agreement of the observed and calculated results. In column 4 of Table I, the ratio of the scattering power of the vapour saturated at $t°$ to the scattering power of the vapour saturated at 35°, and in column 5 the ratio of the vapour densities under the same conditions, taken from Ramsay and Young's data, are shown. If Rayleigh's law of scattering had been valid throughout this range, the figures in columns 4 and 5 should have agreed. Actually there is an

Table I.—Scattering by Saturated Ether Vapour.

Temperature ($t°$ C.).	Scattering at $t°$ C. / Scattering by liquid ether at 35° C.		Ratio of scattering at $t°$ C to scattering at 35° C.	Ratio of density of vapour at $t°$ C. to density at 35° C.
	Observed.	Calculated from Einstein-Smoluchowski formula.		
33	0·038	0·036	1·00	0·92
54	0·075	0·074	1·97	1·81
75	0·150	0·140	3·95	3·22
91	0·22	0·23	5·79	4·79
108	0·45	0·45	11·8	7·1
123·5	0·71	0·71	18·7	10·0
144·5	1·31	1·29	34·5	15·7
164	2·7	3·0	71·0	24·0
170·5	3·9	4·0	103	28·0
179·2	9·1	7·9	239	35·3
183·5	13·0	12·2	342	39·8
186·5	21	19	553	44·2
190	42	39	1105	51·9
191	68	55	1658	53·8

enormous difference, showing that, at least in the case of ether, Rayleigh's law breaks down completely, and that the "principle of random phase," on which he founded his theory of molecular scattering, is inapplicable except at very low pressures, where the deviation from Boyle's law is negligible. The scattering power of the vapour is greater, and at the higher temperatures enormously greater, than if the molecules of the gas scattered secondary waves in arbitrary phase-relationships with each other.

The experimental results of the scattering power of ether in the gaseous phase above the critical point are shown in the second column of Table II. In this case, in order to calculate the compressibilities and refractive indices, it is necessary to know the actual mass of ether used, and the volume of the bulb in which it is contained, and these were separately determined at the completion of the experiment, and were found to be 3·164 grm. and 13·98 c.c. respectively. The theoretical value of the scattering power as given by the Einstein-Smoluchowski formula is shown in column 3.

Table II.—Scattering by Ether above Critical Temperature.

Temperature (t° C.).	$\dfrac{\text{Scattering at } t\text{° C.}}{\text{Scattering by liquid ether at 35° C.}}$		Ratio of scattering at t° C. to scattering at 35° C.	Ratio of density of vapour at t° C. to density at 35° C.
	Observed.	Calculated from Einstein-Smoluchowski formula.		
194	378	200 at 195°	9950	71·6
196	172	200 at 197°	4530	71·6
198	126			
200	84	87	3316	71·6
202	64	65	1680	71·6
212·5	27	30	710	71·6
217	21	26	553	71·6

Considering the unavoidable experimental uncertainties such as a slight fluctuation of temperature, which has an enormous influence in the neighbourhood of the critical point, the presence of a trace of dissolved air or other impurity in the ether, and the difficulty of determining the value of the compressibility accurately, the agreement should be considered good. The figures in columns 4 and 5 further illustrate the fact that the scattering power of the gas throughout the range of temperature is far greater than would be the case if the "principle of random phase" were valid.

Finally, Table III shows the observed scattering power of liquid ether.

The density data for liquid ether have been taken from Ramsay and Young's paper; the compressibilities below 100° from Amagat's determina-

Table III.—Scattering by Liquid Ether.

Temperature (t° C.).	$\frac{\text{Scattering at } t^\circ \text{ C.}}{\text{Scattering at } 35^\circ \text{ C.}}$.	
	Observed.	Calculated from Einstein-Smoluchowski formula.
33	1·00	1·00
61	1·23	1·34
80	1·85	1·65
91	1·90	1·83
104	2·3	2·13 at 100°
125	3·0	
130	3·8	
145	4·9	8·2 at 150°
153	6·3	
170	12·0	14·1 at 175°
179	22	
185	38	27·2
190·5	82	73 at 190°

tions, and those above 100° from Ramsay and Young's work. There seems to be an inaccuracy in Ramsay and Young's compressibility data for liquid ether at 150°; the values of the pressure and volume do not plot into a smooth curve, and in the curve accompanying Ramsay and Young's paper one can notice an unusually rapid change of curvature in the close proximity of the vapourisation point. Making allowance for this, a satisfactory agreement appears between the theoretical and observed values of the scattering power.

Taking the results contained in the three Tables together, it will be seen that the scattering rises rapidly as the critical point is approached, and the values for the saturated vapour, liquid and gaseous phases approach each other and converge to a very large value at this point. The intensity of the scattering under critical conditions is enormous, being about 750 times that of liquid ether at ordinary temperature.

4. *Critical Opalescence.*

The Einstein-Smoluchowski formula (1), as it stands, gives an infinite value for the scattering power at the critical point. This is physically impossible, and both Smoluchowski and Einstein have pointed out that, in the immediate neighbourhood of the critical point, terms involving $(d^3p/dv^3)_t$ and higher order terms should be taken into account. It is, however, doubtful if even the inclusion of these terms would be sufficient, for the scattered light is markedly less blue in this region, which Einstein's theory does not contemplate. Ornstein and Zernike* have pointed out that the independence of density-fluctuations in different elements of volume assumed by Einstein is inadmissible in this region, and, taking these mutual influences into account, they have deduced a modified equation for the opalescence in which the term $-(dp/dv)_t$ is replaced by

$$-\left(\frac{dp}{dv}\right)_t + \frac{4\pi RT}{v^2}\frac{\epsilon^2}{\mu^2\lambda^2}, \tag{3}$$

where v is the molecular volume and ϵ is a measure of the radius of action of a molecule, beyond which its influence is negligible. Using this, the total fraction of light scattered per unit-volume, instead of being infinity at the critical point, has a finite value

$$\frac{2\pi}{27}\frac{\mu^2(\mu^2-1)^2\,(\mu^2+2)^2}{N\lambda^2\epsilon^2},\ v. \tag{4}$$

Also the proportionality of the scattering light with λ^{-4} gives place to one of proportionality with λ^{-2}.

The change of colour of the opalescent light in the neighbourhood of the critical point, is easily observable when the opalescent beam is reduced in intensity by means of the rotating disc photometer to equality with the scattered light from liquid ether. If we assume that the maximum value of the scattering coefficient is given by (4), we can easily obtain an estimate of ϵ. Observations of the maximum intensity of the scattered light were made by heating the ether about 10° above the critical temperature, and then *very slowly* cooling it. With the particular quantity of ether used (3·164 grm. in 13·98 c.c.) the maximum scattering was found to be 750 times that of liquid ether at 35° C. Putting in the numerical values, and taking λ equal to be 5000 Å.U., the scattering coefficient works out to be $8{\cdot}6\times10^{-3}$, and ϵ to be $4{\cdot}6\times10^{-7}$ cm.

However, this correction to the Einstein-Smoluchowski formula is of secondary importance, and does not affect the broad aspects of the question.

5. *Polarisation of the Scattered Light.*

It has been shown by Lord Rayleigh, that the light scattered by dust-free gases and vapours in a direction at right angles to the incident beam is, in most cases, imperfectly polarised. In the case of ether vapour, he found that the ratio of the weak component to the strong in the scattered light was 1·7 per cent. It was interesting to study how the imperfection of polarisation changes as the condition of the substance changes from liquid to vapour. At ordinary temperatures, the light scattered by liquid ether showed an imperfection of polarisation amounting to 8 per cent. As the temperature was raised, it remained constant until a temperature of about 125° was reached and then fell off, rapidly at first, and then more slowly to the value corresponding to that of the vapour at the critical point. Table IV gives the results of the polarisation measurements on liquid ether at different temperatures.

---

* 'Roy. Soc. Proc. Amsterdam,' vol. 17 (1914), and 'Phys. Zeitschrift,' vol. 18 (1917).

Table IV.—Polarisation of Light Scattered by Liquid Ether.

Temperature.	Weak component compared with strong.
33° C.	8 per cent.
60	8 „
80	7 „
100	7·5 „
120	7·5 „
135	5 „
145	3 „
165	1·7 „
189	1·3 „

In the case of the vapour, measurements of polarisation were made from 160° to about 200° C. The values of the ratio of the weak component to the strong ranged from 1·3 to 1·0 per cent. There was no change in the value of the ratio on passing through the critical point.

The difference between the liquid and the vapour at the lower temperatures is striking. In the case of gases, the late Lord Rayleigh has shown that imperfection of polarisation can be explained on the assumption of asymmetry of the molecules. The increased value for the imperfect polarisation in liquids, and its rapid decrease with temperature, are results of great interest which await explanation.* Further studies on this point, with a number of other vapours and liquids, are being made by the writer.

This imperfection of polarisation has to be taken account of in calculating the intensity of the scattered light. In the case of gases, Cabannes has shown that the scattering is greater in the ratio $6(1+r)/(6-7r)$, where r is the ratio of the weak component to the strong in the light scattered in a perpendicular direction. The question can be viewed in yet another way. The imperfectly polarised light can be looked upon as being caused by an admixture of a certain quantity of unpolarised light with the polarised light. If to a beam of polarised light of intensity A, there be added a beam of unpolarised light of intensity 2B, the resultant light will be imperfectly polarised and the ratio of the weak component to the strong would be given by $r = \mathrm{B}/(\mathrm{A}+\mathrm{B})$ and the intensity of the resultant light will be $\mathrm{A}(1+2\mathrm{B}/\mathrm{A}) = \mathrm{A}\left(1+\frac{2r}{1-r}\right)$. When r is small this reduces to $\mathrm{A}(1+2r)$. An expression of this form is found to agree with the intensity of the scattered light in gases and liquids, when A is taken to be the scattering calculated from the Einstein-Smoluchowski formula. The factor $1+2r/(1-r)$ becomes of great importance when r is large. It will be noticed that when r is not very large, this expression becomes practically equal to the correction factor $6(1+r)/(6-7r)$ introduced by Cabannes. In the Tables I, II and III given above such a correction has been introduced. The imperfection of polarisation for ether vapour has been throughout taken to be 1·2 per cent. and that of liquid has been taken as given in Table IV.

---

* Since the above was written, a theory based on the conception of molecular æolotropy has been developed by Prof. C. V. Raman, which not only explains these phenomena, but predicts also that the imperfectness of polarisation, both in the case of liquids and vapours, should decrease and tend to a minimum at the critical point. Observations by the writer on benzene confirm this result.

6. *Summary.*

1. Three instances of light scattering by homogeneous media are known; opalescence of a substance near the critical point, scattering of light by gases, and scattering of light by liquids. The present investigation was undertaken to test the view put forward by Prof. C. V. Raman that these phenomena are fundamentally related to each other, and that the intensity of scattering in all these cases should be represented by the Einstein-Smoluchowski formula, which is more general than the Rayleigh law of scattering and supersedes it except in the special case of gases obeying Boyle's law.

2. Experiments on the scattering of light by ether, in the vapour and liquid phases at different temperatures from 33° C. up to the critical temperature 193·6°, and in the gaseous phase from 193·6° to 217°, are described.

3. The results are in accord with the Einstein-Smoluchowski formula and not with the Rayleigh law of scattering.

4. The Einstein-Smoluchowski formula is inapplicable in the *immediate neighbourhood* of the critical point. The scattered light is markedly less blue here. Following the theoretical work of Ornstein and Zernike, from the maximum value of the intensity of the scattered light, the value of ϵ the radius of action of an ether molecule is deduced to be $4{\cdot}6 \times 10^{-7}$ cm.

5. The light scattered at right angles to the incident beam is imperfectly polarised; the ratio of the weak component to the strong is throughout nearly 1·2 per cent. in the case of vapour, while, in the case of liquids, the

ratio is 8 per cent. at ordinary temperatures which remains constant till about 120°, and then falls off, rapidly at first and then slowly to about 1·2 per cent. at the critical point. There is no change of imperfection of polarisation on passing through the critical point. The correction due to this in the expression for the intensity of the scattered light is given.

In conclusion, I have great pleasure in recording my indebtedness to Prof. C. V. Raman who suggested the research, and who continued to take an inspiring interest in its progress. The experimental work was carried out in the Physical Laboratory of the Indian Association for the Cultivation of Science, Calcutta.

The Motion of Ellipsoidal Particles Immersed in a Viscous Fluid.

By G. B. Jeffery, M.A., D.Sc., Fellow of University College, London.

(Communicated by Prof. L. N. G. Filon, F.R.S. Received January 30, 1922.)

§ 1. *Introduction.*

In both physical and biological science, we are often concerned with the properties of a fluid, or plasma, in which small particles or corpuscles are suspended and carried about by the motion of the fluid. The presence of the particles will influence the properties of the suspension in bulk, and, in particular, its viscosity will be increased. The most complete mathematical treatment of the problem, from this point of view, has been that given by Einstein,[*] who considered the case of spherical particles and gave a simple formula for the increase in the viscosity. We have extended this work to the case of particles of ellipsoidal shape.

The second section of the paper is occupied with the requisite solution of the equations of motion of the fluid. The problem of the motion of a viscous fluid, due to an ellipsoid moving through it with a small velocity of translation in a direction parallel to one of its axes, has been solved by Oberbeck,[†] and the corresponding problem for an ellipsoid rotating about one of its axes by Edwards.[‡] In both cases the equations of motion are approximated by neglecting the terms involving the squares of the velocities. It may be seen,

---

[*] "Eine neue Bestimmung der Moleküldimensionen," 'Ann. d. Physik,' vol. 19, p. 289 (1896); with a correction in vol. 34, p. 591 (1911).

[†] 'Crelle,' vol. 81 (1876).

[‡] 'Quart. Jour. Math.,' vol. 26 (1892).

a posteriori, that the condition for the validity of this approximation is that the product of the velocity of the ellipsoid by its linear dimensions shall be small compared with the "kinematic coefficient of viscosity" of the fluid. In relation to our present problem, it will therefore be satisfied either for sufficiently slow motions, *or* for sufficiently small particles.

The third section is devoted to the consideration of the forces acting upon the particle as a whole. It is found that these reduce to two couples, one tending to make the particle adopt the same rotation as the surrounding fluid, and the other tending to set the particle with its axes parallel to the principal axes of distortion of the surrounding fluid.

In the fourth section we investigate the motion of a particle which is subject to no forces except those arising from the pressure of the fluid on its surface. It is found that there is a variety of possible motions corresponding to different initial conditions. They are periodic and of a general character somewhat like the motion of a top. There is thus a certain degree of indeterminacy in the problem as treated by the approximation which neglects the squares of the velocities.

Sections five and six are concerned with the dissipation of energy and with the viscosity of suspensions. It is found that the increase of viscosity may still be represented by a formula of Einstein's type with a modified numerical factor. Owing to the indeterminacy revealed in section four, it is not possible to give a definite value for this numerical factor. It is, however, possible to specify an upper and a lower limit for this factor, and the difference between these limits diminishes to zero as the particles approach a spherical shape.

Section seven is devoted to a tentative suggestion for the removal of this indeterminacy. It is, that when a variety of motions is possible for given boundary conditions under the approximated equations of motion, the actual motion will be that which corresponds to minimum dissipation of energy. If this hypothesis be granted, it is possible to state many of the results of the paper with greater precision.

§ 2. *The Solution of the Equations of Motion.*

Consider, in the first instance, the modification produced in the motion of a viscous fluid by the presence of a single ellipsoidal particle. We shall assume that, apart from the disturbance produced in the immediate neighbourhood of the particle, the motion is steady, and varies in space on a scale which is large compared with the dimensions of the particle. It appears from the work of Oberbeck that, if the fluid at a distance from the particle has a velocity of translation relative to the particle, the latter is acted upon by a resultant force. It will appear in the course of the investigation that any other motion of the fluid can only give rise to a resultant couple. The particle will therefore ultimately assume the velocity of translation appropriate to that part of the fluid which it displaces. On the above assumptions this velocity of translation will be sensibly uniform, and we may therefore take the centre of the particle as at rest.

Let x, y, z be rectangular co-ordinates referred to axes fixed in the particle and moving with it, and let the surface of the particle be the ellipsoid

$$x^2/a^2+y^2/b^2+z^2/c^2=1. \tag{1}$$

The undisturbed motion of the fluid in the neighbourhood of the particle is then given by

$$\left.\begin{aligned} u_0 &= \mathbf{a}x+\mathbf{h}y+\mathbf{g}z+\eta z-\zeta y \\ v_0 &= \mathbf{h}x+\mathbf{b}y+\mathbf{f}z+\zeta x-\xi y \\ w_0 &= \mathbf{g}x+\mathbf{f}y+\mathbf{c}z+\xi y-\eta x \end{aligned}\right\}, \tag{2}$$

where $\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$, $\mathbf{f}$, $\mathbf{g}$, $\mathbf{h}$, ξ, η, ζ are the components of distortion and rotation of the fluid. These are taken to be constant in space through a volume which is large compared with the dimensions of the particle. They will, however, vary with the time, since the particle will rotate under the influence of the fluid, and the axes of x, y, z will be rotating axes.

Let the spins of the ellipsoid about its axes be ω_1, ω_2, ω_3. Take a set of axes, x', y', z', in fixed directions with their origin at the centre of the particle, and let the components of distortion and rotation of the fluid in its undisturbed motion referred to these axes be $\mathbf{a}'$, $\mathbf{b}'$, $\mathbf{c}'$, ... These are constant, but the corresponding unaccented quantities are linear functions of these, with coefficients which depend upon the direction cosines of the two sets of axes, and which therefore vary with the time in consequence of ω_1, ω_2, ω_3.

Neglecting squares and products of the velocities, the equations of motion of an incompressible viscous fluid referred to moving axes are of the type

$$\mu\nabla^2u-\frac{\partial p}{\partial x}=\rho\left\{\frac{\partial u}{\partial t}-\omega_3 v+\omega_2 w\right\}, \tag{3}$$

where ρ, μ, p are respectively the density, coefficient of viscosity, and mean pressure of the fluid.

The spins ω_1, ω_2, ω_3 are produced by the motion of the fluid, and, as will appear explicitly later, they are of the same order as the velocities. Hence to our order of approximation all products like $\omega_3 v$ may be neglected. It is not quite so obvious that the remaining terms on the right-hand side of (3) are of the second order. It will appear later that u, v, w are homogeneous linear functions of $\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$, ..., with constant coefficients. They are, therefore, homogeneous linear functions of the constants $\mathbf{a}'$, $\mathbf{b}'$, $\mathbf{c}'$, ..., with coefficients which are functions of the direction cosines of the moving axes. Hence the only terms which can arise in $\partial u/\partial t$ are of the type $\omega_1 a'$, which is of the second order of small quantities, and is therefore to be neglected.

The equations of motion therefore reduce to

$$\mu\nabla^2 u = \frac{\partial p}{\partial x}, \qquad \mu\nabla^2 v = \frac{\partial p}{\partial y}, \qquad \mu\nabla^2 w = \frac{\partial p}{\partial z}, \tag{4}$$

with the equation of continuity

$$\frac{\partial u}{\partial x}+\frac{\partial v}{\partial y}+\frac{\partial w}{\partial z} = 0. \tag{5}$$

We have to find a solution of (4) and (5) which agrees with (2) at great distances from the origin, and which gives on the surface of the ellipsoid

$$u = \omega_2 z-\omega_3 y, \qquad v = \omega_3 x-\omega_1 z, \qquad w = \omega_1 y-\omega_2 x. \tag{6}$$

As usual, let λ denote the positive root of

$$\frac{x^2}{a^2+\lambda}+\frac{y^2}{b^2+\lambda}+\frac{z^2}{c^2+\lambda} = 1, \tag{7}$$

and

$$\Delta = \{(a^2+\lambda)(b^2+\lambda)(c^2+\lambda)\}^{\frac{1}{2}}, \tag{8}$$

and let

$$\alpha = \int_\lambda^\infty \frac{d\lambda}{(a^2+\lambda)\Delta}, \qquad \beta = \int_\lambda^\infty \frac{d\lambda}{(b^2+\lambda)\Delta}, \qquad \gamma = \int_\lambda^\infty \frac{d\lambda}{(c^2+\lambda)\Delta}. \tag{9}$$

We will also write

$$\alpha' = \int_\lambda^\infty \frac{d\lambda}{(b^2+\lambda)(c^2+\lambda)\Delta}, \qquad \alpha'' = \int_\lambda^\infty \frac{\lambda d\lambda}{(b^2+\lambda)(c^2+\lambda)\Delta}, \tag{10}$$

with symmetrical integrals for β', β'', γ', γ'', so that

$$\alpha' = \frac{\gamma-\beta}{b^2-c^2}, \qquad \alpha'' = \frac{b^2\beta-c^2\gamma}{b^2-c^2}. \tag{11}$$

We will denote the corresponding integrals, in which the lower limit of integration has been replaced by zero, by α_0, β_0', ...

We have the following well-known solutions of Laplace's equation :—

$$\Omega = \int_\lambda^\infty \left\{\frac{x^2}{a^2+\lambda}+\frac{y^2}{b^2+\lambda}+\frac{z^2}{c^2+\lambda}-1\right\}\frac{d\lambda}{\Delta} \tag{12}$$

and

$$\chi_1 = \alpha' yz, \qquad \chi_2 = \beta' zx, \qquad \chi_3 = \gamma' xy. \tag{13}$$

We shall seek a solution in which u, v, w are expressed as linear functions of the first and second differential coefficients of Ω and the first differential coefficients of χ_1, χ_2, χ_3. These latter can, of course, be expressed in terms of the second differential coefficients of Ω, but as the corresponding terms enter into the solution in a somewhat different way, it will be convenient to keep them separate.

From (7) we have

$$\frac{\partial\lambda}{\partial x}=\frac{2x\mathrm{P}^2}{a^2+\lambda},\qquad \frac{\partial\lambda}{\partial y}=\frac{2y\mathrm{P}^2}{b^2+\lambda},\qquad \frac{\partial\lambda}{\partial z}=\frac{2z\mathrm{P}^2}{c^2+\lambda} \tag{14}$$

where

$$\frac{1}{\mathrm{P}^2}=\frac{x^2}{(a^2+\lambda)^2}+\frac{y^2}{(b^2+\lambda)^2}+\frac{z^2}{(c^2+\lambda)^2}. \tag{15}$$

We then have a number of formulæ of which the following are typical:

$$\frac{\partial\Omega}{\partial x}=2\alpha x,\qquad \frac{\partial^2\Omega}{\partial x^2}=2\alpha-\frac{4x^2\mathrm{P}^2}{(a^2+\lambda)^2\Delta},\qquad \frac{\partial^2\Omega}{\partial y\,\partial z}=-\frac{4yz\mathrm{P}^2}{(b^2+\lambda)(c^2+\lambda)\Delta}, \tag{16}$$

and

$$\frac{\partial\chi_1}{\partial x}=\frac{\partial\chi_2}{\partial y}=\frac{\partial\chi_3}{\partial z}=-\frac{2xyz\mathrm{P}^2}{\Delta^3},$$

$$\frac{\partial\chi_1}{\partial y}=\alpha' z-\frac{2y^2z\mathrm{P}^2}{(b^2+\lambda)^2(c^2+\lambda)\Delta},\qquad \frac{\partial\chi_1}{\partial z}=\alpha' y-\frac{2yz^2\mathrm{P}^2}{(b^2+\lambda)(c^2+\lambda)^2\Delta}. \tag{17}$$

All these differential coefficients tend to zero at infinity.

We assume

$$\begin{aligned} u=u_0&+\frac{\partial}{\partial x}(\mathrm{R}\chi_1+\mathrm{S}\chi_2+\mathrm{T}\chi_3)+\mathrm{W}\frac{\partial\chi_3}{\partial y}-\mathrm{V}\frac{\partial\chi_2}{\partial z}\\ &+\mathrm{A}\left(x\frac{\partial^2\Omega}{\partial x^2}-\frac{\partial\Omega}{\partial x}\right)+\mathrm{H}\left(x\frac{\partial^2\Omega}{\partial x\,\partial y}-\frac{\partial\Omega}{\partial y}\right)+\mathrm{G}'\left(x\frac{\partial^2\Omega}{\partial x\,\partial z}-\frac{\partial\Omega}{\partial z}\right)\\ &+y\left(\mathrm{H}'\frac{\partial^2\Omega}{\partial x^2}+\mathrm{B}\frac{\partial^2\Omega}{\partial x\,\partial y}+\mathrm{F}\frac{\partial^2\Omega}{\partial x\,\partial z}\right)\\ &+z\left(\mathrm{G}\frac{\partial^2\Omega}{\partial x^2}+\mathrm{F}'\frac{\partial^2\Omega}{\partial x\,\partial y}+\mathrm{C}\frac{\partial^2\Omega}{\partial x\,\partial z}\right), \end{aligned} \tag{18}$$

$$\begin{aligned} v=v_0&+\frac{\partial}{\partial y}(\mathrm{R}\chi_1+\mathrm{S}\chi_2+\mathrm{T}\chi_3)+\mathrm{U}\frac{\partial\chi_1}{\partial z}-\mathrm{W}\frac{\partial\chi_3}{\partial x}\\ &+x\left(\mathrm{A}\frac{\partial^2\Omega}{\partial x\,\partial y}+\mathrm{H}\frac{\partial^2\Omega}{\partial y^2}+\mathrm{G}'\frac{\partial^2\Omega}{\partial y\,\partial z}\right)\\ &+\mathrm{H}'\left(y\frac{\partial^2\Omega}{\partial x\,\partial y}-\frac{\partial\Omega}{\partial x}\right)+\mathrm{B}\left(y\frac{\partial^2\Omega}{\partial y^2}-\frac{\partial\Omega}{\partial y}\right)+\mathrm{F}\left(y\frac{\partial^2\Omega}{\partial y\,\partial z}-\frac{\partial\Omega}{\partial z}\right)\\ &+z\left(\mathrm{G}\frac{\partial^2\Omega}{\partial x\,\partial y}+\mathrm{F}'\frac{\partial^2\Omega}{\partial y^2}+\mathrm{C}\frac{\partial^2\Omega}{\partial y\,\partial z}\right), \end{aligned} \tag{19}$$

and

$$\begin{aligned} w=w_0&+\frac{\partial}{\partial z}(\mathrm{R}\chi_1+\mathrm{S}\chi_2+\mathrm{T}\chi_3)+\mathrm{V}\frac{\partial\chi_2}{\partial x}-\mathrm{U}\frac{\partial\chi_1}{\partial y}\\ &+x\left(\mathrm{A}\frac{\partial^2\Omega}{\partial x\,\partial z}+\mathrm{H}\frac{\partial^2\Omega}{\partial y\,\partial z}+\mathrm{G}'\frac{\partial^2\Omega}{\partial z^2}\right)\\ &+y\left(\mathrm{H}'\frac{\partial^2\Omega}{\partial x\,\partial z}+\mathrm{B}\frac{\partial^2\Omega}{\partial y\,\partial z}+\mathrm{F}\frac{\partial^2\Omega}{\partial z^2}\right)\\ &+\mathrm{G}\left(z\frac{\partial^2\Omega}{\partial x\,\partial z}-\frac{\partial\Omega}{\partial x}\right)+\mathrm{F}'\left(z\frac{\partial^2\Omega}{\partial y\,\partial z}-\frac{\partial\Omega}{\partial y}\right)+\mathrm{C}\left(z\frac{\partial^2\Omega}{\partial z^2}-\frac{\partial\Omega}{\partial z}\right). \end{aligned} \tag{20}$$

It may be seen by inspection that these values of u, v, w satisfy (5), since $\mathbf{a}+\mathbf{b}+\mathbf{c}=0$ for an incompressible fluid, and Ω and χ_1, χ_2, χ_3 satisfy Laplace's equation.

Substituting these expressions into (4), it is easily seen that these are satisfied provided that

$$p = p_0+2\mu\left\{\mathrm{A}\frac{\partial^2\Omega}{\partial x^2}+\mathrm{B}\frac{\partial^2\Omega}{\partial y^2}+\mathrm{C}\frac{\partial^2\Omega}{\partial z^2}+(\mathrm{F}+\mathrm{F}')\frac{\partial^2\Omega}{\partial y\,\partial z}\right.$$
$$\left.+(\mathrm{G}+\mathrm{G}')\frac{\partial^2\Omega}{\partial z\,\partial x}+(\mathrm{H}+\mathrm{H}')\frac{\partial^2\Omega}{\partial x\,\partial y}\right\}, \tag{21}$$

where p_0 is the constant mean pressure at a distance from the ellipsoid. It therefore remains to satisfy the conditions (6) on the surface of the ellipsoid $\lambda = 0$. Substituting from (16) and (17) into (18), (19), (20), we have

$$\begin{aligned}u = {} & x\{\mathbf{a}+\gamma'\mathrm{W}-\beta'\mathrm{V}-2(\alpha+\beta+\gamma)\mathrm{A}\}\\
& +y\{\mathbf{h}-\zeta+\gamma'\mathrm{T}-2\beta\mathrm{H}+2\alpha\mathrm{H}'\}\\
& +z\{\mathbf{g}+\eta+\beta'\mathrm{S}-2\gamma\mathrm{G}'+2\alpha\mathrm{G}\}\\
& -\frac{2x\mathrm{P}^2}{(a^2+\lambda)\Delta}\Big[\{\mathrm{R}+2(b^2+\lambda)\mathrm{F}+2(c^2+\lambda)\mathrm{F}'\}yz/(b_2+\lambda)(c^2+\lambda)\\
& \qquad +\{\mathrm{S}+2(c^2+\lambda)\mathrm{G}+2(a^2+\lambda)\mathrm{G}'\}zx/(c^2+\lambda)(a^2+\lambda)\\
& \qquad +\{\mathrm{T}+2(a^2+\lambda)\mathrm{H}+2(b^2+\lambda)\mathrm{H}'\}xy/(a^2+\lambda)(b^2+\lambda)\\
& \qquad +\{\mathrm{W}-2(a^2+\lambda)\mathrm{A}+2(b^2+\lambda)\mathrm{B}\}y^2/(b^2+\lambda)^2\\
& \qquad -\{\mathrm{V}-2(c^2+\lambda)\mathrm{C}+2(a^2+\lambda)\mathrm{A}\}z^2/(c^2+\lambda)^2\Big],\end{aligned} \tag{22}$$

$$\begin{aligned}v = {} & x\{\mathbf{h}+\zeta+\gamma'\mathrm{T}+2\beta\mathrm{H}-2\alpha\mathrm{H}'\}\\
& +y\{\mathbf{b}+\alpha'\mathrm{U}-\gamma'\mathrm{W}-2(\alpha+\beta+\gamma)\mathrm{B}\}\\
& +z\{\mathbf{f}-\xi+\alpha'\mathrm{R}-2\gamma\mathrm{F}+2\beta\mathrm{F}'\}\\
& -\frac{2y\mathrm{P}^2}{(b^2+\lambda)\Delta}\Big[\{\mathrm{R}+2(b^2+\lambda)\mathrm{F}+2(c^2+\lambda)\mathrm{F}'\}yz/(b^2+\lambda)(c^2+\lambda)\\
& \qquad +\{\mathrm{S}+2(c^2+\lambda)\mathrm{G}+2(a^2+\lambda)\mathrm{G}'\}zx/(c^2+\lambda)(a^2+\lambda)\\
& \qquad +\{\mathrm{T}+2(a^2+\lambda)\mathrm{H}+2(b^2+\lambda)\mathrm{H}'\}xy/(a^2+\lambda)(b^2+\lambda)\\
& \qquad +\{\mathrm{U}-2(b^2+\lambda)\mathrm{B}+2(c^2+\lambda)\mathrm{C}\}z^2/(c^2+\lambda)^2\\
& \qquad -\{\mathrm{W}-2(a^2+\lambda)\mathrm{A}+2(b^2+\lambda)\mathrm{B}\}x^2/(a^2+\lambda)^2\Big],\end{aligned} \tag{23}$$

and

$$
\begin{aligned}
w = {} & x\{\mathbf{g}-\eta+\beta'\mathrm{S}-2\alpha\mathrm{G}+2\gamma\mathrm{G}'\} \\
& +y\{\mathbf{f}+\xi+\alpha'\mathrm{R}+2\gamma\mathrm{F}-2\beta\mathrm{F}'\} \\
& +z\{\mathbf{c}+\beta'\mathrm{V}-\alpha'\mathrm{U}-2(a+\beta+\gamma)\mathrm{C}\} \\
& -\frac{2z\mathrm{P}^2}{(c^2+\lambda)\Delta}\Big[\{\mathrm{R}+2(b^2+\lambda)\mathrm{F}+2(c^2+\lambda)\mathrm{F}'\}yz/(b^2+\lambda)(c^2+\lambda) \\
& \qquad +\{\mathrm{S}+2(c^2+\lambda)\mathrm{G}+2(a^2+\lambda)\mathrm{G}'\}zx/(c^2+\lambda)(a^2+\lambda) \\
& \qquad +\{\mathrm{T}+2(a^2+\lambda)\mathrm{H}+2(b^2+\lambda)\mathrm{H}'\}xy/(a^2+\lambda)(b^2+\lambda) \\
& \qquad +\{\mathrm{V}-2(c^2+\lambda)\mathrm{C}+2(a^2+\lambda)\mathrm{A}\}x^2/(a^2+\lambda)^2 \\
& \qquad -\{\mathrm{U}-2(b^2+\lambda)\mathrm{B}+2(c^2+\lambda)\mathrm{C}\}y^2/(b^2+\lambda)^2\Big]. \qquad (24)
\end{aligned}
$$

Putting $\lambda = 0$, identifying with (6), and equating coefficients, we obtain equations which determine the constants. On solution we find their values to be as follows :—

$$
\left.\begin{aligned}
\mathrm{A} &= \tfrac{1}{6}\{2\alpha_0''\mathbf{a}-\beta_0''\mathbf{b}-\gamma_0''\mathbf{c}\}/(\beta_0''\gamma_0''+\gamma_0''\alpha_0''+\alpha_0''\beta_0'') \\
\mathrm{B} &= \tfrac{1}{6}\{2\beta_0''\mathbf{b}-\gamma_0''\mathbf{c}-\alpha_0''\mathbf{a}\}/(\beta^{0\prime\prime}\gamma_0''+\gamma_0''\alpha_0''+\alpha_0''\beta_0'') \\
\mathrm{C} &= \tfrac{1}{6}\{2\gamma_0''\mathbf{c}-\alpha_0''\mathbf{a}-\beta_0''\mathbf{b}\}/(\beta_0''\gamma_0''+\gamma_0''\alpha_0''+\alpha_0''\beta_0'')
\end{aligned}\right\}, \qquad (25)
$$

$$
\left.\begin{aligned}
\mathrm{F} &= \frac{\beta_0\mathbf{f}-c^2\alpha_0(\xi-\omega_1)}{2\alpha_0'(b^2\beta_0+c^2\gamma_0)}, & \mathrm{F}' &= \frac{\gamma_0\mathbf{f}+b^2\alpha_0(\xi-\omega_1)}{2\alpha_0'(b^2\beta_0+c^2\gamma_0)} \\
\mathrm{G} &= \frac{\gamma_0\mathbf{g}-a^2\beta_0(\eta-\omega_2)}{2\beta_0'(c^2\gamma_0+a^2\alpha_0)}, & \mathrm{G}' &= \frac{\alpha_0\mathbf{g}+c^2\beta_0(\eta-\omega_2)}{2\beta_0'(c^2\gamma_0+a^2\alpha_0)} \\
\mathrm{H} &= \frac{\alpha_0\mathbf{h}-b^2\gamma_0(\zeta-\omega_3)}{2\gamma_0'(a^2\alpha_0+b^2\beta_0)}, & \mathrm{H}' &= \frac{\beta_0\mathbf{h}+a^2\gamma_0(\zeta-\omega_3)}{2\gamma_0'(a^2\alpha_0+b^2\beta_0)}
\end{aligned}\right\}, \qquad (26)
$$

$$
\mathrm{R} = -\mathbf{f}/\alpha_0', \qquad \mathrm{S} = -\mathbf{g}/\beta_0', \qquad \mathrm{T} = -\mathbf{h}/\gamma_0'. \qquad (27)
$$

The values of U, V, W will not be needed for our purpose. They are given by

$$
\left.\begin{aligned}
\mathrm{U} &= 2b^2\mathrm{B}-2c^2\mathrm{C} \\
\mathrm{V} &= 2c^2\mathrm{C}-2a^2\mathrm{A} \\
\mathrm{W} &= 2a^2\mathrm{A}-2b^2\mathrm{B}
\end{aligned}\right\}. \qquad (28)
$$

Thus the constants in (18), (19), and (20) are all uniquely determined and the velocity of the fluid is known at all points.

For later use we note the following relations :—

$$
\mathrm{A}+\mathrm{B}+\mathrm{C} = 0, \qquad (29)
$$

$$
\mathrm{U}+\mathrm{V}+\mathrm{W} = 0, \qquad (30)
$$

and, using the relation $\mathbf{a}+\mathbf{b}+\mathbf{c} = 0$,

$$
\mathrm{A}\mathbf{a}+\mathrm{B}\mathbf{b}+\mathrm{C}\mathbf{c} = \tfrac{1}{2}(\alpha_0''\mathbf{a}^2+\beta_0''\mathbf{b}^2+\gamma_0''\mathbf{c}^2)/(\beta_0''\gamma_0''+\gamma_0''\alpha_0''+\alpha_0''\beta_0''). \qquad (31)
$$

§ 3. *The Resultant Couple on the Particle.*

For an incompressible fluid the stresses are given, with the usual notation, by

$$\widehat{xx} = -p+2\mu\frac{\partial u}{\partial x}, \qquad \widehat{yz} = \mu\left(\frac{\partial w}{\partial y}+\frac{\partial v}{\partial z}\right). \tag{32}$$

The calculation of the values of these stresses on the surface of the ellipsoid is greatly simplified by noting that most of the brackets in (22), (23), and (24) vanish for $\lambda = 0$, and that, therefore, the only terms surviving in the surface values of the differential coefficients of the velocity components are those arising from the differentiation of the brackets.

If we denote the components of force per unit area acting on the surface of the ellipsoid by X, Y, Z, we have

$$\left.\begin{aligned} X &= \{\widehat{xx}\,x/a^2+\widehat{xy}\,y/b^2+\widehat{xz}\,z/c^2\} \\ Y &= \{\widehat{yx}\,x/a^2+\widehat{yy}\,y/b^2+\widehat{yz}\,z/c^2\} \\ Z &= \{\widehat{zx}\,x/a^2+\widehat{zy}\,y/b^2+\widehat{zz}\,z/c^2\} \end{aligned}\right\}. \tag{33}$$

The reduction is straightforward and, after simplifying as far as possible by means of the relations between the constants, we obtain

$$\begin{aligned} X &= -p_0\mathrm{P}\,\frac{x}{a^2}+\frac{8\mu\mathrm{P}}{abc}\left\{\mathrm{A}\,\frac{x}{a^2}+\mathrm{H}\,\frac{y}{b^2}+\mathrm{G}'\,\frac{z}{c^2}\right\}-4\mu\mathrm{P}(\alpha_0\mathrm{A}+\beta_0\mathrm{B}+\gamma_0\mathrm{C})\,\frac{x}{a^2}, \\ Y &= -p_0\mathrm{P}\,\frac{y}{b^2}+\frac{8\mu\mathrm{P}}{abc}\left\{\mathrm{H}'\,\frac{x}{a^2}+\mathrm{B}\,\frac{y}{b^2}+\mathrm{F}\,\frac{z}{c^2}\right\}-4\mu\mathrm{P}(\alpha_0\mathrm{A}+\beta_0\mathrm{B}+\gamma_0\mathrm{C})\,\frac{y}{b^2}, \\ Z &= -p_0\mathrm{P}\,\frac{z}{c^2}+\frac{8\mu\mathrm{P}}{abc}\left\{\mathrm{G}\,\frac{x}{a^2}+\mathrm{F}'\,\frac{y}{b^2}+\mathrm{C}\,\frac{z}{c^2}\right\}-4\mu\mathrm{P}(\alpha_0\mathrm{A}+\beta_0\mathrm{B}+\gamma_0\mathrm{C})\,\frac{z}{c^2}. \end{aligned} \tag{34}$$

Since, apart from P, which is even in x, y, z, only first powers of the co-ordinates appear in these expressions, they will separately vanish on integration over the surface of the ellipsoid. Hence there is no resultant *force* acting on the particle. Denoting the components of the resultant couple acting on the particle by L, M, N, we have on integration over the surface of the ellipsoid

$$\mathrm{L} = \iint (y\mathrm{Z}-z\mathrm{Y})\,d\mathrm{S},$$

which on substitution from (34), omitting terms which obviously vanish on integration, gives

$$\mathrm{L} = \frac{8\mu}{abc}\iint\left(\mathrm{F}'\,\frac{y^2}{b^2}-\mathrm{F}\,\frac{z^2}{c^2}\right)d\mathrm{S}.$$

This integral is readily evaluated in terms of the volume of the ellipsoid, and we have from this and two similar expressions

$$L = \tfrac{32}{3}\pi\mu(F'-F), \quad M = \tfrac{32}{3}\pi\mu(G'-G), \quad N = \tfrac{32}{3}\pi\mu(H'-H). \tag{35}$$

Substituting from (26), these give

$$\left.\begin{aligned} L &= \frac{16\pi\mu}{3(b^2\beta_0+c^2\gamma_0)}\{(b^2-c^2)\mathbf{f}+(b^2+c^2)(\xi-\omega_1)\}, \\ M &= \frac{16\pi\mu}{3(c^2\gamma_0+a^2\alpha_0)}\{(c^2-a^2)\mathbf{g}+(c^2+a^2)(\eta-\omega_2)\}, \\ N &= \frac{16\pi\mu}{3(a^2\alpha_0+b^2\beta_0)}\{(a^2-b^2)\mathbf{h}+(a^2+b^2)(\zeta-\omega_3)\}. \end{aligned}\right\}. \tag{36}$$

The resultant couple acting upon the ellipsoid may thus be regarded as consisting of two parts: one which vanishes when the ellipsoid has the same resultant spin as the fluid, and one which vanishes when the axes of the ellipsoid coincide with the principal axes of distortion of the fluid.

§ 4. *The Motion of the Particle—Laminar Motion.*

If the particle is subject to no forces except those exerted by the fluid upon its surface, then in the slow motions to which our assumptions have already restricted us, the resultant couple on the particle must vanish at every instant. Hence, from (36),

$$\left.\begin{aligned} (b^2+c^2)\omega_1 &= b^2(\xi+\mathbf{f})+c^2(\xi-\mathbf{f}) \\ (c^2+a^2)\omega_2 &= c^2(\eta+\mathbf{g})+a^2(\eta-\mathbf{g}) \\ (a^2+b^2)\omega_3 &= a^2(\zeta+\mathbf{h})+b^2(\zeta-\mathbf{h}) \end{aligned}\right\}, \tag{37}$$

or, from (35),

$$F' = F, \qquad G' = G, \qquad H' = H. \tag{38}$$

These relations will simplify many of our formulæ and, in particular, we may note that (26) are replaced by

$$F = \frac{\mathbf{f}}{2\alpha_0'(b^2+c^2)}, \qquad G = \frac{\mathbf{g}}{2\beta_0'(c^2+a^2)}, \qquad H = \frac{\mathbf{h}}{2\gamma_0'(a^2+b^2)}. \tag{39}$$

It has not been found possible to obtain a solution of equations (37) which would give the motion of the particle for the most general possible undisturbed motion of the fluid. Some light may, however, be thrown on the problem by the consideration of a particular, but important, case.

Take axes x', y', z', fixed in direction, and let the components of the fluid velocity in the undisturbed motion be u_0', v_0', w_0', parallel to these axes. Consider the "laminar motion" given by $u_0' = v_0' = 0$, $w_0' = \kappa y'$, where κ is a constant. Let the direction cosines of the axes of the ellipsoid (*i.e.*, the axes x, y, z) referred to the axes of x', y', z' be (l_1, m_1, n_1), (l_2, m_2, n_2) and

(l_3, m_3, n_3) respectively. Then employing the usual formulæ for the transformation of axes, we obtain

$$\begin{aligned}
\mathbf{a} &= \kappa m_1 n_1, & \mathbf{b} &= \kappa m_2 n_2, & \mathbf{c} &= \kappa m_3 n_3, \\
\mathbf{f} &= \tfrac{1}{2}\kappa(m_2 n_3 + m_3 n_2), & \mathbf{g} &= \tfrac{1}{2}\kappa(m_3 n_1 + m_1 n_3), & \mathbf{h} &= \tfrac{1}{2}\kappa(m_1 n_2 + m_2 n_1), \\
\xi &= \tfrac{1}{2}\kappa(m_2 n_3 - m_3 n_2), & \eta &= \tfrac{1}{2}\kappa(m_3 n_1 - m_1 n_3), & \zeta &= \tfrac{1}{2}\kappa(m_1 n_2 - m_2 n_1),
\end{aligned} \quad (40)$$

and equations (37) become

$$\left.\begin{aligned}
(b^2+c^2)\,\omega_1 &= \kappa(b^2 m_2 n_3 - c^2 m_3 n_2) \\
(c^2+a^2)\,\omega_2 &= \kappa(c^2 m_3 n_1 - a^2 m_1 n_3) \\
(a^2+b^2)\,\omega_3 &= \kappa(a^2 m_1 n_2 - b^2 m_2 n_1)
\end{aligned}\right\}. \quad (41)$$

These may be regarded as the equations of motion of the ellipsoid and give its spins for any position of its axes. It may be seen from these that a possible type of motion is that in which any one of the three axes of the ellipsoid lies permanently along the axis of x', *i.e.*, perpendicular both to the direction of the velocity and to that of the velocity gradient in the undisturbed motion, while the ellipsoid rotates about this axis with a variable spin. A typical motion of this class is that for which

$$\begin{aligned}
&l_1, m_1, n_1 = 1, 0, 0\,; \qquad l_2, m_2, n_2 = 0, \cos\chi, \sin\chi\,; \\
&l_3, m_3, n_3 = 0, -\sin\chi, \cos\chi\,; \qquad \omega_1 = \dot\chi\,; \qquad \omega_2 = \omega_3 = 0.
\end{aligned} \quad (42)$$

Equations (41) then give

$$(b^2+c^2)\,\dot\chi = \kappa(b^2\cos^2\chi + c^2\sin^2\chi),$$

or, omitting a constant of integration which is obviously immaterial,

$$\tan\chi = \frac{b}{c}\tan\frac{bc\kappa t}{b^2+c^2}. \quad (43)$$

The equations readily admit of integration in the case of an ellipsoid of revolution. Let $b = c$, and introduce the three Euler angles, θ the angle between the axes of x and x', ϕ the angle between the planes of $x'y'$ and $x'x$, ψ the angle between the planes of $x'x$ and xy. Then

$$\omega_1 = \dot\phi\cos\theta + \dot\psi, \quad \omega_2 = \dot\theta\sin\psi - \dot\phi\sin\theta\cos\psi, \quad \omega_3 = \dot\theta\cos\psi + \dot\phi\sin\theta\sin\psi,$$

and

$$\left.\begin{aligned}
m_1 &= \sin\theta\cos\phi, \qquad n_1 = \sin\theta\sin\phi \\
m_2 &= -\sin\phi\sin\psi + \cos\theta\cos\phi\cos\psi \\
n_2 &= \cos\phi\sin\psi + \cos\theta\sin\phi\cos\psi \\
m_3 &= -\sin\phi\cos\psi - \cos\theta\cos\phi\sin\psi \\
n_3 &= \cos\phi\cos\psi - \cos\theta\sin\phi\sin\psi
\end{aligned}\right\}. \quad (44)$$

Substituting in (41), we have after some reduction,

$$\dot{\phi}\cos\theta + \dot{\psi} = \tfrac{1}{2}\kappa\cos\theta, \tag{45}$$

$$(a^2+b^2)\dot{\theta} = \kappa(a^2-b^2)\sin\theta\cos\theta\sin\phi\cos\phi, \tag{46}$$

$$(a^2+b^2)\dot{\phi} = \kappa(a^2\cos^2\phi + b^2\sin^2\phi). \tag{47}$$

From (47) we have, if $t = 0$ when $\phi = 0$,

$$\tan\phi = \frac{a}{b}\tan\frac{\kappa abt}{a^2+b^2}, \tag{48}$$

and dividing (46) by (47) and integrating,

$$\tan^2\theta = \frac{a^2b^2}{k^2(a^2\cos^2\phi + b^2\sin^2\phi)}, \tag{49}$$

where k is a constant of integration.

It appears that the motion of the ellipsoid is periodic. Its axis of revolution describes a cone about the perpendicular to the plane of the undisturbed motion, to which its inclination varies between the limits $\tan^{-1}(a/k)$ and $\tan^{-1}(b/k)$.

This statement requires modification in the degenerate cases of a disc and a thin rod respectively. Putting $a = 0$ in equations (45), (46), and (47), we find that a disc can take any position in which its faces are entirely composed of stream-lines of the undisturbed motion of the fluid. It thus moves through the fluid edge-on with its axis at any inclination to the fluid shear. If we put $b = c = 0$ in the same equations we find what is otherwise obvious, that a thin rod can set itself in any direction in a plane which is a plane of constant velocity for the undisturbed motion of the fluid. In what follows, however, we shall regard rods and discs as spheroids of small but finite thickness.

This investigation reveals no tendency on the part of the ellipsoid to set its axis in any particular direction with regard to the undisturbed motion of the fluid. The motions corresponding to all values of k are consistent with the boundary conditions and with the approximated equations of motion.

§ 5. *The Dissipation of Energy.*

As the velocity of the fluid is finite at great distances from the ellipsoid, the total rate of dissipation of energy will be infinite. We will accordingly calculate the excess of the rate of dissipation of energy over that which obtains in the same motion of the fluid in the absence of the ellipsoidal particle. Difficulties which arise in connection with the boundary conditions, make it imperative that we should define precisely the two fluid motions which we are thus comparing.

Describe a sphere S, of large but finite radius R, whose centre is at the centre of the particle. Let the velocity of the fluid at points on S be given by equations (2), and let the ellipsoid move freely under the fluid pressures on its surface so that relations (38) are satisfied. We will calculate the excess of the rate of dissipation of energy over that which would occur if the ellipsoid were removed, the sphere S completely filled with fluid, and the velocity at all points on its surface the same as in the first case.

The direct calculation of the rate of dissipation of energy presents very great difficulties. For our purpose, it will be sufficient to consider the rate at which work must be done over the boundary of the sphere S in order to maintain the motion. This, of course, differs from the rate of dissipation of energy by the rate of increase of the kinetic energy of the fluid. For the periodic motions investigated in the last section, however, the average rate of work of maintaining the motion must be the same as the average rate of dissipation of energy.

As we shall ultimately suppose the radius of the sphere S to tend to infinity, we may adopt such a degree of approximation that negative powers of R are neglected in the final result. This implies that the stresses must be correct to order R^{-3} and the velocities to order R^{-2}.

At great distances from the origin we have, neglecting terms of order r^{-3},

$$\Omega = -\tfrac{4}{3}\, 1/r, \qquad \chi_1 = \chi_2 = \chi_3 = 0. \tag{50}$$

Substituting in (18), (19), (20), and remembering that (38) are satisfied, and writing

$$\Phi \equiv Ax^2 + By^2 + Cz^2 + 2Fyz + 2Gzx + 2Hxy, \tag{51}$$

we have, neglecting terms of order r^{-4},

$$u = u_0 - 4x\Phi/r^5, \qquad v = v_0 - 4y\Phi/r^5, \qquad w = w_0 - 4z\Phi/r^5, \tag{52}$$

while (21) gives

$$p = p_0 - 8\mu\Phi/r^5. \tag{53}$$

Superpose a motion, finite at all points within the sphere S, and such as to cancel the second terms in the right-hand members of equations (52) at points on its surface. Such a motion is easily found by the ordinary spherical harmonic methods, and when combined with the motion represented by (52), we have,

$$\left.\begin{aligned} u &= u_0 - \frac{4x(R^5 - r^5)}{R^5 r^5}\Phi + \frac{5(R^2 - r^2)}{R^5}\frac{\partial\Phi}{\partial x} \\ v &= v_0 - \frac{4y(R^5 - r^5)}{R^5 r^5}\Phi + \frac{5(R^2 - r^2)}{R^5}\frac{\partial\Phi}{\partial y} \\ w &= w_0 - \frac{4z(R^5 - r^5)}{R^5 r^5}\Phi + \frac{5(R^2 - r^2)}{R^5}\frac{\partial\Phi}{\partial z} \end{aligned}\right\}, \tag{54}$$

which give on the surface of S

$$p = p_0 - 50\mu\Phi/R^5. \tag{55}$$

These velocities will not accurately satisfy the boundary conditions on the surface of the ellipsoid, but proceeding by successive approximations it is easily seen that the additional terms will be of order R^{-7} on the surface of S, and may therefore be neglected. The stresses at $r = R$ are calculated from (54) and (55), by the aid of (32). Typical results are

$$\widehat{xx} = -p_0 + 2\mu\mathbf{a} + 10\mu\left\{\frac{5}{R^5}\Phi + \frac{4x^2}{R^7}\Phi - \frac{2x}{R^5}\frac{\partial\Phi}{\partial x}\right\}$$

$$\widehat{yz} = 2\mu\mathbf{f} + 10\mu\left\{\frac{4yz}{R^7}\Phi - \frac{z}{R^5}\frac{\partial\Phi}{\partial y} - \frac{y}{R^5}\frac{\partial\Phi}{\partial z}\right\}. \tag{56}$$

From these we have

$$\widehat{xr} = (x\,\widehat{xx} + y\,\widehat{xy} + z\,\widehat{xz}/r)R$$

$$= -\frac{x}{R}p_0 + \frac{2\mu}{R}(\mathbf{a}x + \mathbf{h}y + \mathbf{g}z) + 10\mu\left\{\frac{7x}{R^6}\Phi - \frac{1}{R^4}\frac{\partial\Phi}{\partial x}\right\}, \tag{57}$$

and two similar equations.

The rate at which work is done in maintaining the motion is given by

$$\iint (u_0\,\widehat{xr} + v_0\,\widehat{yr} + w_0\,\widehat{zr})\,dS,$$

where the integration extends over the surface of S. On substitution from (57), this integral is readily evaluated and it gives

$$\tfrac{4}{3}\pi R^3 \,.\, 2\mu(\mathbf{a}^2 + \mathbf{b}^2 + \mathbf{c}^2 + 2\mathbf{f}^2 + 2\mathbf{g}^2 + 2\mathbf{h}^2)$$

$$+ \tfrac{32}{3}\pi\mu(\mathbf{a}A + \mathbf{b}B + \mathbf{c}C + 2\mathbf{f}F + 2\mathbf{g}G + 2\mathbf{h}H). \tag{58}$$

The first term is the work necessary to maintain the motion in the absence of the particle; the second represents the addition due to its presence. On making the radius of the sphere S tend to infinity, this result becomes exact. On substitution from (31) and (39) this gives for the additional rate of doing work owing to the presence of the particle

$$\frac{dW}{dt} = \tfrac{16}{3}\pi\mu\left\{\frac{\alpha_0''\mathbf{a}^2 + \beta_0''\mathbf{b}^2 + \gamma_0''\mathbf{c}^2}{\beta_0''\gamma_0'' + \gamma_0''\alpha_0'' + \alpha_0''\beta_0''}\right.$$

$$\left. + \frac{2\mathbf{f}^2}{\alpha_0'(b^2 + c^2)} + \frac{2\mathbf{g}^2}{\beta_0'(c^2 + a^2)} + \frac{2\mathbf{h}^2}{\gamma_0'(a^2 + b^2)}\right\}. \tag{59}$$

It is interesting to note that if due regard is not paid to the conditions on the bounding sphere, and if the calculation is made directly from equations (52), the result is only one-fifth of that given above.

In the case of an ellipsoid of revolution ($b = c$) we have γ_0', γ_0'' respectively equal to β_0', β_0'', and it is easy to show that $\beta_0'' + 2\alpha_0'' = 2b^2\alpha_0'$. Equation (59) then takes the simplified form

$$\frac{dW}{dt} = \tfrac{16}{3}\pi\mu\left\{\frac{\alpha_0''\mathbf{a}^2}{2b^2\alpha_0'\beta_0''} + \frac{\mathbf{b}^2 + \mathbf{c}^2 + 2\mathbf{f}^2}{2b^2\alpha_0'} + \frac{2(\mathbf{g}^2 + \mathbf{h}^2)}{\beta_0'(a^2 + b^2)}\right\}. \tag{60}$$

If the fluid is moving in laminar motion, this gives on substitution from (40) and (44)

$$\frac{dW}{dt} = \tfrac{4}{3}\pi\mu\kappa^2\left\{\left(\frac{\alpha_0''}{2b^2\alpha_0'\beta_0''}+\frac{1}{2b^2\alpha_0'}-\frac{2}{\beta_0'(a^2+b^2)}\right)\sin^4\theta\sin^2 2\phi\right.$$
$$\left.+\frac{1}{b^2\alpha_0'}\cos^2\theta+\frac{2}{\beta_0'(a^2+b^2)}\sin^2\theta\right\}. \qquad (61)$$

The average values of $\cos^2\theta$ and $\sin^4\theta\sin^2 2\phi$ are calculated from (48) and (49) as follows:

$$\frac{\kappa ab}{2\pi(a^2+b^2)}\int_0^{2\pi}\cos^2\theta\frac{dt}{d\phi}d\phi = \frac{k^2}{\sqrt{\{(a^2+k^2)(b^2+k^2)\}}}$$

and

$$\frac{\kappa ab}{2\pi(a^2+b^2)}\int_0^{2\pi}\sin^4\theta\sin^2 2\phi\frac{dt}{d\phi}d\phi = \frac{2a^2b^2}{(a^2-b^2)^2}\left\{\frac{a^2+b^2+2k^2}{\sqrt{\{(a^2+k^2)(b^2+k^2)\}}}-2\right\}.$$

If D is the average rate of dissipation of energy, (61) then gives

$$D = \tfrac{4}{3}\pi\mu\kappa^2\left[\frac{2a^2b^2}{(a^2-b^2)^2}\left\{\frac{a^2+b^2+2k^2}{\sqrt{\{(a^2+k^2)(b^2+k^2)\}}}-2\right\}\right.$$
$$\left\{\frac{\alpha_0''}{2b^2\alpha_0'\beta_0''}+\frac{1}{2b^2\alpha_0'}-\frac{2}{\beta_0'(a^2+b^2)}\right\}$$
$$\left.+\frac{k^2}{\sqrt{\{(a^2+k^2)(b^2+k^2)\}}}\left\{\frac{1}{b^2\alpha_0'}-\frac{2}{\beta_0'(a^2+b^2)}\right\}+\frac{2}{\beta_0'(a^2+b^2)}\right]. \qquad (62)$$

We will show that this has no maxima or minima for finite, non-zero values of k. Differentiating with regard to k^2, we obtain

$$\frac{dD}{d(k^2)} = -\frac{2\pi\mu\kappa^2[k^2\beta_0''\{2b^2\alpha_0'-(a^2+b^2)\beta_0'\}+a^2b^2\beta_0'(\alpha_0''-\beta_0'')]}{3b^2\alpha_0'\beta_0'\beta_0''\{(a^2+k^2)(b^2+k_2)\}^{3/2}}. \qquad (63)$$

From their definition in equations (10), we see that α_0', α_0'', β_0', β_0'' are all positive, while $\alpha_0''-\beta_0''$ is of the same sign as a^2-b^2. Further

$$2b^2\alpha_0'-(a^2+b^2)\beta_0' = (a^2-b^2)\int_0^\infty\frac{(b^2-\lambda)\,d\lambda}{(a^2+\lambda)^{3/2}(b^2+\lambda)^3}$$
$$= \tfrac{3}{2}(a^2-b^2)\int_0^\infty\frac{\lambda d\lambda}{(a^2+\lambda)^{5/2}(b^2+\lambda)^2},$$

after an integration by parts. The latter integral is clearly positive, and hence the left-hand side is of the same sign as a^2-b^2. It follows that the right-hand side of (63) never vanishes for finite values of k, and is of opposite sign to a^2-b^2. Accordingly, the motion which gives minimum average dissipation of energy will correspond to $k=\infty$ for a prolate spheroid, and to $k=0$ for an oblate spheroid. The average rate of dissipation of energy will, in all cases, lie between the values corresponding to $k=0$ and $k=\infty$.

We will write

$$D = \nu v\mu\kappa^2 \qquad (64)$$

where ν is a factor and $v = 4\pi ab^2/3$, and is therefore the volume of the particle. Equation (62) then gives, – for $k = \infty$,

$$\nu = 1/ab^4\alpha_0', \tag{65}$$

and for $k = 0$

$$\nu = \frac{1}{ab^3(a+b)^2}\left\{\frac{ab^2}{\beta_0''}+\frac{a}{2\alpha_0'}+\frac{2b}{\beta_0'}\right\}. \tag{66}$$

In the case of spheroids the integrals in (9) and (10) can be evaluated in finite terms. We obtain,

(1) for prolate spheroids, $b = c = a\cos\theta$

$$\begin{aligned} a^5\alpha_0' &= \tfrac{1}{4}\operatorname{cosec}^4\theta\,\{(2-5\cos^2\theta)\sec^4\theta+3\Theta\}, \\ a^5\beta_0' &= \operatorname{cosec}^4\theta\,\{2+\sec^2\theta-3\Theta\}, \\ a^3\beta_0'' &= \operatorname{cosec}^4\theta\,\{(2+\cos^2\theta)\Theta-3\}, \end{aligned} \tag{67}$$

where $\Theta \equiv \operatorname{cosec}\theta\log\tan\left(\dfrac{\theta}{2}+\dfrac{\pi}{4}\right)$,

(2) for oblate spheroids, $b = c = a\sec\theta$,

$$\begin{aligned} b^5\alpha_0' &= \tfrac{1}{4}\operatorname{cosec}^4\theta\,\{3\Theta-\cos\theta\,(5-2\cos^2\theta)\}, \\ b^5\beta_0' &= \operatorname{cosec}^4\theta\,\{(2+\cos^2\theta)\sec\theta-3\Theta\}, \\ b^3\beta_0'' &= \operatorname{cosec}^4\theta\,\{(1+2\cos^2\theta)\Theta-3\cos\theta\}, \end{aligned} \tag{68}$$

where $\Theta \equiv \theta/\sin\theta$.

The maximum and minimum values of ν for spheroids of different shapes are given in Tables I and II. They are given in terms of the "ellipticity" of the meridian section of the particle, as this is generally more convenient for measurement than the eccentricity. The ellipticity ϵ is defined to be the difference of the greatest and least diameters divided by the greatest, so that $\epsilon = (a-b)/a$ for a prolate spheroid and $\epsilon = (b-a)/b$ for an oblate spheroid.

Table I.—Prolate Spheroids.

Ellipticity $\epsilon = \dfrac{a-b}{a}$.	Minimum ν.	Maximum ν.
0	2·5	2·5
0·1	2·431	2·540
0·2	2·361	2·586
0·3	2·295	2·645
0·4	2·232	2·719
0·5	2·174	2·819
0·6	2·120	2·958
0·7	2·073	3·170
0·8	2·035	3·548
0·9	2·010	4·485
1·0	2·000	∞

Table II.—Oblate Spheroids.

Ellipticity $\epsilon = \dfrac{b-a}{b}$.	Minimum ν.	Maximum ν.
0	2·5	2·5
0·1	2·464	2·582
0·2	2·426	2·683
0·3	2·388	2·818
0·4	2·348	3·003
0·5	2·306	3·267
0·6	2·262	3·670
0·7	2·216	4·354
0·8	2·168	5·744
0·9	2·116	9·960
1·0	2·061	∞

It will be noted that the maximum value of ν tends to infinity in each case as the ellipticity tends to unity. This does not, of course, mean that the dissipation of energy becomes infinite, since the volume of the particle at the same time tends to zero.

§ 6. *The Viscosity of Suspensions.*

The results of the last section may be applied, as Einstein has applied the corresponding results for a sphere, to the determination of the viscosity of a fluid containing solid spheroidal particles in suspension. If the suspension is so dilute that the distances between neighbouring particles are great compared with their dimensions, we may regard each particle as the cause of a certain increase in the dissipation of energy.

In the case of spherical particles the motion is steady and D is given accurately, without averaging, by the second term in (58). It is then easy to show from (10), (25) and (39), that, when $a = b = c$,

$$\mathbf{A}/\mathbf{a} = \mathbf{B}/\mathbf{b} = \mathbf{C}/\mathbf{c} = \mathbf{F}/\mathbf{f} = \mathbf{G}/\mathbf{g} = \mathbf{H}/\mathbf{h} = 5a^3/8. \tag{69}$$

If v is the volume of the particle, so that $v = 4\pi a^3/3$, we have,

$$D = 5v\mu\,(\mathbf{a}^2+\mathbf{b}^2+\mathbf{c}^2+2\mathbf{f}^2+2\mathbf{g}^2+2\mathbf{h}^2).$$

Hence, if v is the total volume of the particles in unit volume of the suspension, the total rate of dissipation of energy is

$$2\mu\,(1+2{\cdot}5\,\mathrm{V})\,(\mathbf{a}^2+\mathbf{b}^2+\mathbf{c}^2+2\mathbf{f}^2+2\mathbf{g}^2+2\mathbf{h}^2).$$

From the point of view of the dissipation of energy, we may, therefore, regard the suspension as a homogeneous fluid whose coefficient of viscosity is μ^*, where

$$\mu^* = \mu\,(1+2{\cdot}5\,\mathrm{V}). \tag{70}$$

This agrees with Einstein's well-known result. It does not require the particles to be equal in size, so long as they are all spherical in shape. It applies only to suspensions which are sufficiently dilute and in which there is no tendency to form aggregates. With these restrictions it is true for any steady motion of the suspension, subject only to the condition that, if $\mathbf{a}, \mathbf{b}, \ldots$ vary in space, they shall do so only on a scale which is large compared with the dimensions of the particles, and that this variation shall not promote a variation in the concentration of the particles. For example, in applying this formula to measurements of viscosity by capillary tube methods, we should have to be assured: (1) that the diameter of the tube is large compared with that of the particles; (2) that there is no tendency for the particles to select particular paths, say, along the axis of the tube.

Some further light may be thrown upon this last point. It is easily seen that, within the limitations of the present theory, there is no resultant force on the particle in a direction perpendicular to the axis of the tube. If the axis of the tube is taken as axis of z, the undisturbed motion is given by $u = v = 0, w = A(a^2 - x^2 - y^2)$, where a is the radius of the capillary tube and A is a known constant. Referred to parallel axes through the centre of a particle (h, k), this gives $u = v = 0, w = a^2 - h^2 - k^2 - x^2 - y^2 - 2hx - 2ky$. This may be regarded as the superposition of two motions: (1) $u = v = 0$, $w = a^2 - h^2 - k^2 - x^2 - y^2$; (2) $u = v = 0, w = -2hx - 2ky$. Neither of these motions can give rise to any force on the particle perpendicular to the axis of z—the motion (1), for reasons of symmetry, and the motion (2), by the investigations of the present paper. It follows that the combined motion can give rise to no such force.

In the case of laminar motion, we can extend Einstein's result to ellipsoidal particles. We have then from (64) for the average total dissipation of energy per unit volume of the suspension

$$\mu\kappa^2(1+\nu V)$$

giving in place of (70)

$$\mu^* = \mu(1+\nu V). \tag{71}$$

The values of ν for prolate and oblate spheroids of different ellipticity of meridian section are given in Tables I and II respectively.

It will be noted that the minimum value of ν is always less than the value (2·5) for spherical particles. In this connection it may be remarked that, although in some cases Einstein's formula agrees well with experiment, in others it agrees only if ν has a value greater than 2·5.* Smoluchowski† proposes a formula of the type (71) for non-spherical particles, and, with regard to the factor ν, he remarks: "For the sphere this value ($\nu = 2·5$) is evidently a minimum, since by virtue of its rolling motion it influences the shearing motion of the least volume of the fluid." This argument is not valid, for we have shown that for any spheroid there are possible motions for which the increase in the dissipation of energy is less than it is for a sphere of the same volume.

§ 7. *A Hypothesis on the Dissipation of Energy.*

It appears from the foregoing paragraphs that we are able only to specify limits for the numerical factor in Einstein's formula. It is obviously undesirable to leave a problem, which is physically quite determinate, in this indeterminate form.

---

* For an account of the many experimental investigations of this problem, see Arrhenius, 'Biochem. Jour.,' vol. 11, p. 112 (1917).

† "Theoretische Bemerkungen über die Viskosität der Kolloide," 'Kolloid Zeitsch.,' vol. 18, p. 191 (1916).

We might assume that at any instant the axes of the particles are in random directions. On this hypothesis it is easy to calculate the rate of dissipation of energy. It is found that this gives a value for Einstein's factor which is greater than the maximum values shown in Tables I and II. The explanation is that random directions for the axes of the particles do not correspond to a permanent steady state.

Again, we can determine ν if we know the number of particles for which the k of (49) lies between k and $k+dk$, say $f(k)\,dk$. This, however, is just what our present theory fails to reveal: according to it, all values of k are possible, and there is no indication of their relative probability. It would appear that this failure is due to the limitations of the theory of the slow motion of a viscous fluid.

It will be remembered that the only stable direction of permanent translation of an ellipsoid in a *perfect* fluid is that of its least axis. If an attempt is made to solve the corresponding problem for a viscous fluid by Overbeck's method, it is found that no couple is exerted upon the ellipsoid by the fluid if its motion is one of pure translation in any direction whatever. A heavy ellipsoid falling through a viscous fluid would fall with different velocities for different directions of its axes with regard to the vertical, but apparently it would not tend to set its axes in any particular directions. This is due to the fact that in the case of the perfect fluid the resultant couples arise from those terms in the pressure which depend upon the square of the velocity, terms which are neglected in the ordinary viscous theory.

It seems not improbable that, in our present problem, a more complete investigation would reveal the fact that the particles do tend to adopt special orientations with respect to the motion of the surrounding fluid. Our suggestion is that *the particles will tend to adopt that motion which, of all the motions possible under the approximated equations, corresponds to the least dissipation of energy*. If this be true, it reconciles the perfect fluid theory, the viscous theory, and experiment in the case of an ellipsoid falling under gravity through a fluid.

Our assumption is reminiscent of the theorems of Helmholtz and Korteweg on the rate of dissipation of energy in the motion of a viscous fluid. A little reflection, however, will show that these theorems are not really applicable to our present problem. For one thing, except in the case of spheroidal particles, no steady motion is possible. Further, these theorems are themselves deduced from the approximate equations of motion obtained by neglecting the squares and products of the velocities, and our investigations have been based upon exact solutions of these same approximate equations. It is impossible that a general theorem, based upon a set of differential equations, should give, with regard to a particular case, information which is not contained in the corresponding exact solution of the equations.

If this hypothesis be granted, we are able to assert the following propositions:—

1. Prolate spheroidal particles immersed in a fluid in laminar motion will tend to set themselves with their axes perpendicular to the plane of the undisturbed motion of the fluid. They will rotate about their axes with constant angular velocity, and the whole motion will be steady.

2. Oblate spheroidal particles immersed in a fluid in laminar motion will tend to set themselves with an equatorial diameter perpendicular to the plane of the undisturbed motion of the fluid. They will rotate about this diameter with a variable angular velocity, and the whole motion will be periodic but not steady.

3. For motion in a capillary tube, the particles will tend to concentrate along the axis of the tube.

4. The effective value of the numerical factor in Einstein's formula will be the minimum ν, as shown in Tables I and II. This is always less than the 2·5 appropriate to spherical particles. Thus, for particles of given volume, the increase of the viscosity is greatest when the particles are spherical.

An experimental investigation of some of these provisional conclusions would be valuable.

The Motion of a Sphere in a Rotating Liquid.

By G. I. TAYLOR, F.R.S.

(Received June 9, 1922.)

In some recent papers[*] the author has drawn attention to certain general properties of rotating fluids, especially to the differences which may be expected between two- and three-dimensional motion. Unfortunately, mathematical difficulties have so far prevented the solution of any three-dimensional problem in a rotationally moving fluid from being obtained, except in one case, when the motion is very slow. In this case, Prof. Proudman has shown how it is possible to approximate to the solution of the problem of the slow motion of a sphere in a rotating fluid.[†] Even in this case the analysis is very complicated.

There seems little prospect of obtaining a more general solution of the problem when the inertia terms which Proudman neglected are taken into account. On the other hand, it is shown in the following pages that a solution can be obtained in the case when the sphere moves *steadily* along the axis of rotation of the fluid. The limitation imposed by considering only a steady motion necessarily excludes the case considered by Proudman, for all slow steady motions of a rotating fluid are two-dimensional.

In the case when the sphere moves steadily with velocity U along the axis of a fluid rotating with angular velocity Ω, it is possible to reduce the flow to a steady motion by superposing a velocity $-\mathrm{U}$ on the whole system. Since the motion is symmetrical about the axis, only two independent co-ordinates specifying position are necessary, namely, r, the distance of any point from the centre of the sphere, and θ, the angle between the radius from the centre and the axis of symmetry.

Let u be the component of velocity of the fluid along a radius from the centre of the sphere, v the component in an axial plane and perpendicular to the radius, w the component perpendicular to the axial plane. The scheme is shown in fig. 1. Since the motion is symmetrical, the equation of continuity is satisfied if

$$u = -\frac{1}{r^2 \sin\theta}\frac{\partial\psi}{\partial\theta}, \qquad v = \frac{1}{r\sin\theta}\frac{\partial\psi}{\partial r}, \tag{1}$$

where ψ is Stokes' stream function.

---

[*] 'Roy. Soc. Proc.,' A, vol. 100, p. 114 (1921); 'Proc. Camb. Phil. Soc.,' vol. 20, p. 326 (1921); 'Roy. Soc. Proc.,' A, vol. 93, p. 99 (1917).

[†] 'Roy. Soc. Proc.,' A, vol. 92, p. 408 (1916). Proudman's work has recently been extended by Mr. S. F. Grace (these 'Proceedings,' p. 89, *supra*).

Since the motion is symmetrical, the equations of motion are

$$\frac{\mathrm{D}u}{\mathrm{D}t}-\frac{v^2}{r}-\frac{w^2}{r}=-\frac{1}{\rho}\frac{\partial p}{\partial r}, \tag{2}$$

$$\frac{\mathrm{D}v}{\mathrm{D}t}-\frac{w^2\cot\theta}{r}+\frac{uv}{r}=-\frac{1}{\rho r}\frac{\partial p}{\partial \theta}, \tag{3}$$

$$\frac{\mathrm{D}w}{\mathrm{D}t}+\frac{uw}{r}+\frac{vw\cot\theta}{r}=0, \tag{4}$$

where p is the pressure and ρ the density of the fluid.

Since the motion is steady $\frac{\mathrm{D}}{\mathrm{D}t}$ is $u\frac{\partial}{\partial r}+\frac{v}{r}\frac{\partial}{\partial \theta}$.

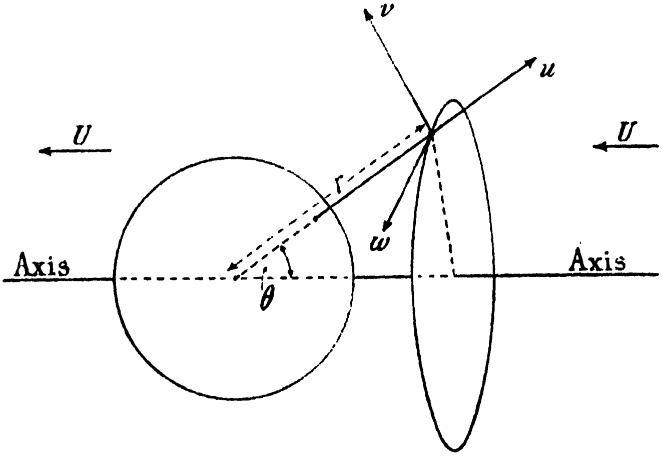


FIG. 1.—Scheme of co-ordinates.

It is easy to verify that $w = \mathrm{A}\psi/(r\sin\theta)$ satisfies the equation (4), and if the motion at infinity consists of a flow with uniform velocity $-\mathrm{U}$ parallel to axis and a rotation about this axis with angular velocity Ω, the constant A is evidently equal to $2\Omega/\mathrm{U}$, so that

$$w = 2\Omega\psi/(\mathrm{U}r\sin\theta). \tag{5}$$

This equation evidently expresses the fact that the circulation in a ring of fluid, which is symmetrical with respect to the axis, remains constant during the motion.

Let us now search for possible solutions of the form

$$\psi = f\sin^2\theta, \tag{6}$$

where f is a function of r only.

The components of velocity are then

$$u = -\frac{2f\cos\theta}{r}, \qquad v = \frac{f'\sin\theta}{r}, \qquad w = \frac{2\Omega}{\mathrm{U}}\frac{f\sin\theta}{r}, \tag{7}$$

where f' is df/dr.

Eliminating p between (2) and (3), it will be found that the form assumed for ψ is legitimate, provided that f satisfies the equation

$$r^3 f''' - 2r^2 f'' - 2rf' + 8f + (4\Omega^2/U^2)(rf' - 2f) = 0. \tag{8}$$

The complete solution of this equation is

$$f = Cz^2 + A\left(\cos z - \frac{\sin z}{z}\right) + B\left(\sin z + \frac{\cos z}{z}\right),$$

or alternatively

$$f = Cz^2 + D\left\{\cos(z+\epsilon) - \frac{\sin(z+\epsilon)}{z}\right\}, \tag{9}$$

where $$k = 2\Omega/U, \qquad z = kr,$$

and A, B, C, D, ϵ are arbitrary constants. When z becomes infinite, the first term in (9) becomes large compared with the others. The motion at a great distance from the sphere is therefore represented by $f = Cz^2$, and, on comparing this with (7), it will be seen that this represents a uniformly rotating fluid moving with velocity $-2Ck^2$ along the axis. Hence

$$C = U/2k^2. \tag{10}$$

The condition at the surface of the sphere, $r = a$, is $u = 0$. Writing $\mu = ka$, this condition gives

$$\tfrac{1}{2}U\frac{\mu^2}{k^2} + D\left\{\cos(\mu+\epsilon) - \frac{\sin(\mu+\epsilon)}{\mu}\right\} = 0. \tag{11}$$

Any values of D and ϵ which satisfy (11) lead to a possible solution of the problem.

It appears, therefore, that there are an infinite number of possible motions round a sphere moving steadily along the axis of a rotating fluid, and that all of them vanish at infinity.

I have not been able to discover how the motion could be set up. Perhaps the different solutions represent the stream-lines due to different ways of starting the motion. It is clear, however, that some of the possible motions represented by (9) could not be set up by starting the sphere from rest; it would be impossible for instance to set up a motion in this way for which f vanished or changed sign anywhere except at the surface of the sphere, because a negative value of f corresponds with a reversed rotation of the fluid about the axis. Such a motion is dynamically possible however, and it corresponds with a case in which all the liquid inside a certain sphere, concentric with the solid sphere, moves with it, forming a kind of sheath of liquid which possesses a rotation about the axis opposite to the rotation at infinity.

It is possible that these considerations may be extended so as to differentiate between the various possible motions corresponding to the various possible pairs of values of D and ϵ which are consistent with (11). The circulation round any symmetrical ring of fluid remains constant during any motion of the fluid. It is equal to $I = 2\pi\Omega y^2_0$, where y_0 is the initial distance of a particle of the ring considered from the axis. If the sphere starts from rest, the total deficiency of fluid in the field which possesses circulation lying between I and $I+(dI/dy_0)\delta y_0$ below that which the fluid would possess in the absence of the sphere, is the volume of fluid displaced by the part of the sphere which is contained between cylinders concentric with the axis, whose radii are y_0 and $y_0+\delta y_0$. This volume is

$$2\pi y_0(a^2-y_0^2)^{\frac{1}{2}}\,\delta y_0 \text{ if } y_0 < a, \quad \text{or} \quad 0 \text{ if } y_0 > a.$$

When the sphere moves in steady motion with velocity U, ψ is connected with y_0 by the equation

$$y_0^2 = 2\psi/U.$$

The total deficiency of fluid possessing circulation lying between I and $I+(dI/dy_0)\delta y_0$ is, therefore, found by calculating the deficiency of fluid possessing circulation between I and $I+(dI/d\psi)\delta\psi$ (where $\delta\psi = Uy_0\delta y_0$), by integrating the total volume of fluid between the stream-lines ψ and $\psi+\delta\psi$ through the whole field. If this deficiency is not equal to $2\pi(a^2-y_0^2)\,\delta y$, when $y_0<a$ and 0 when $y_0>a$, then the motion could not be produced by starting a sphere from rest in a rotating fluid, unless there is a finite motion of the fluid at infinity along the stream-lines close to the axis during the time the motion is being established. This reasoning makes it appear that there is very little chance that any of the motions represented by (9) would be started by moving a sphere from rest in a rotating fluid.

It is interesting to notice that it is possible to find solutions in which $u = v = w = 0$ at the surface of the sphere, so that there is no slipping between the fluid and the surface of the sphere. The condition $U = 0$ at $r = a$ leads to the equation $0 = [f']_{z=\mu}$,

or
$$\frac{\mu U}{k^2}+D\left\{\left(\frac{1}{\mu^2}-1\right)\sin(\mu+\epsilon)-\frac{1}{\mu}\cos(\mu+\epsilon)\right\} = 0. \tag{12}$$

This together with (11) yields the following values for D and ϵ.

$$D = \tfrac{1}{2}U(\mu^4+3\mu^2+9)/k^2, \tag{13}$$

$$\tan(\mu+\epsilon) = 3\mu(3-\mu^2)^{-1}, \tag{14}$$

so that the motion represented by

$$\frac{2k^2}{U}f = z^2+(\mu^4+3\mu^2+9)^{\frac{1}{2}}\left\{\cos(z+\epsilon)-\frac{\sin(z+\epsilon)}{z}\right\}, \tag{15}$$

is one for which the disturbance due to the sphere vanishes at infinity, and it is characterised by the fact that there is no slipping at the surface of the sphere.

This may be a point of some importance because it is the assumption that there is slipping at the surface of a solid body moving in a liquid which vitiates all the ordinary hydrodynamical theories of the motion of solids in fluids. It is possible, therefore, that the solution given above may represent the motion of a sphere in a rotating liquid more closely than the ordinary irrotational solution for a sphere moving in an infinite fluid at rest represents the actual flow in that case.

The surfaces $\psi =$ constant are surfaces of revolution and the stream-lines are spirals wrapped on these surfaces. The sections of the surfaces $\psi =$ constant by an axial plane may be called the stream-lines of the motion in the axial plane. These stream-lines are shown for a particular case in fig. (2). The case chosen is that of a sphere moving along the axis of a rotating fluid at such a speed that it travels a distance equal to its diameter (*i.e.*, $\mu = 2\pi$) during each revolution of the liquid, and the particular solution for which there is no slipping between the sphere and the liquid is chosen. It will be seen that the stream-lines are very different from those which surround a sphere moving in a non-rotating liquid.

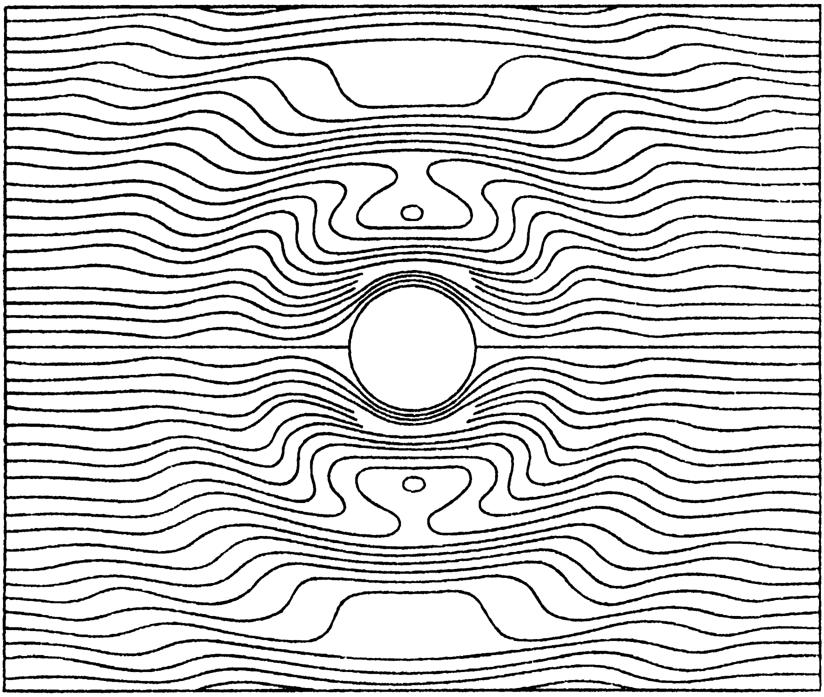


FIG. 2.—Stream-lines due to the motion of a sphere in a rotating fluid. Case when $\mu = 2\pi$.

Propagation of an Isolated Disturbance along the Axis of a Rotating Fluid.

An interesting point about the motion represented by (15) is that it is possible to reduce the radius of the sphere to zero while still retaining a finite velocity in the disturbed motion.

Taking a very small, μ becomes small, and (14) becomes $\tan\epsilon = 0$, so that $\epsilon = 0$, (15) then becomes

$$f = \frac{\mathrm{U}}{2k^2}\left[z^2 + 3\left(\cos z - \frac{\sin z}{z}\right)\right]. \tag{16}$$

It can be verified, by substituting in the original equations of motion, that

$$\psi = \frac{\mathrm{U}}{2k^2}\left[k^2r^2 + 3\left(\cos kr - \frac{\sin kr}{kr}\right)\right]\sin^2\theta, \tag{17}$$

is a solution of the equations of motion, (17) also represents a motion for which the velocity and the pressure are finite and continuous at the origin. The motion consists of a kind of non-rotating core* of liquid propagated with velocity U along the axis of a rotating liquid. It is analogous to the motion produced by a vortex ring, but in an inverse sense.

The stream-lines, due to the motion of a non-rotating core in a rotating fluid, are shown in figs. 3 and 4. The stream-lines of the steady motion, relative to the moving core, are shown in fig. 3. It will be seen that there are no closed stream-lines, so that the disturbance does not carry any fluid with it.

The stream-lines, relative to the main body of the liquid, are shown in fig. 4. It will be seen that the central part of the disturbance resembles Hill's spherical vortex,† and that it is surrounded by spherical waves which travel with it. The analogy between the present disturbance and a spherical vortex is only superficial, for the vortex ring is a mass of rotating fluid which can move through a non-rotating fluid. The present disturbance is a type which could only be propagated in a rotating fluid, and it consists of a core which rotates more slowly than the surrounding fluid and moves parallel to the axis of rotation.

Wave Systems in a Rotating Fluid.

An essential feature of the motions described above is the system of spherical waves which accompany the moving sphere or moving disturbance. That a system of waves would accompany a body moving in a rotating fluid is to be expected. It has been pointed out by Lord Kelvin‡ that rotation confers on a fluid certain properties resembling those of an elastic solid, and in particular, a rotating fluid can transmit waves.

---

* Region on the axis where the rotation about the axis is small compared with the rotation of the fluid as a whole.

† M. J. M. Hill, "On a Spherical Vortex," 'Phil. Trans.,' A, 1894.

‡ Kelvin's 'Mathl. and Physical Papers,' vol. 4, pp. 152, 170.

In order that a system of waves may accompany a disturbance which moves with velocity U along any axis, it is necessary that the velocity of the wave along a normal to the wave-front should be U cos α, where α is the angle between this normal and the direction of the axis along which the disturbance is moving. It is easy to show that a plane wave, of given wave-length, moves in a rotating fluid with velocity proportional to cos α, so that the required condition is satisfied.

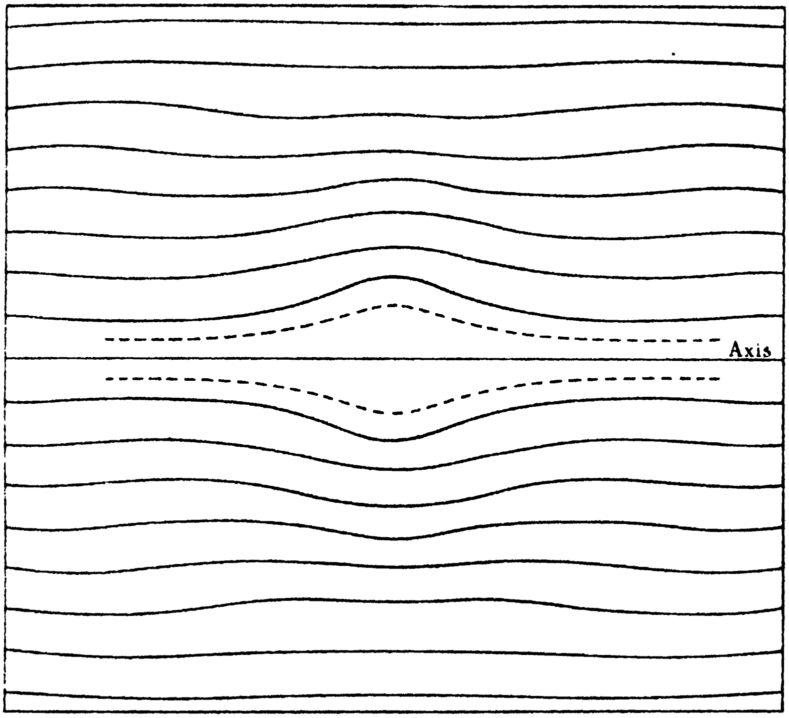


Fig. 3.—Stream-lines relative to the centre of the disturbance due to a non-rotating core travelling along the axis of a rotating fluid.

Writing the equations of motion of a fluid relative to rotating rectangular axes in the form

$$\left.\begin{array}{l} \dfrac{\partial u}{\partial t}+u\dfrac{\partial u}{\partial x}+v\dfrac{\partial u}{\partial y}+w\dfrac{\partial u}{\partial z}-2\Omega u=-\dfrac{\partial}{\partial x}\left(\dfrac{p}{\rho}-\Omega^2r^2\right) \\ \dfrac{\partial v}{\partial t}+u\dfrac{\partial v}{\partial x}+v\dfrac{\partial v}{\partial y}+w\dfrac{\partial v}{\partial z}+2\Omega u=-\dfrac{\partial}{\partial y}\left(\dfrac{p}{\rho}-\Omega^2r^2\right) \\ \dfrac{\partial w}{\partial t}+u\dfrac{\partial w}{\partial x}+v\dfrac{\partial w}{\partial y}+w\dfrac{\partial w}{\partial z}\qquad\quad=-\dfrac{\partial}{\partial z}\left(\dfrac{p}{\rho}-\Omega^2r^2\right) \end{array}\right\}, \qquad (18)$$

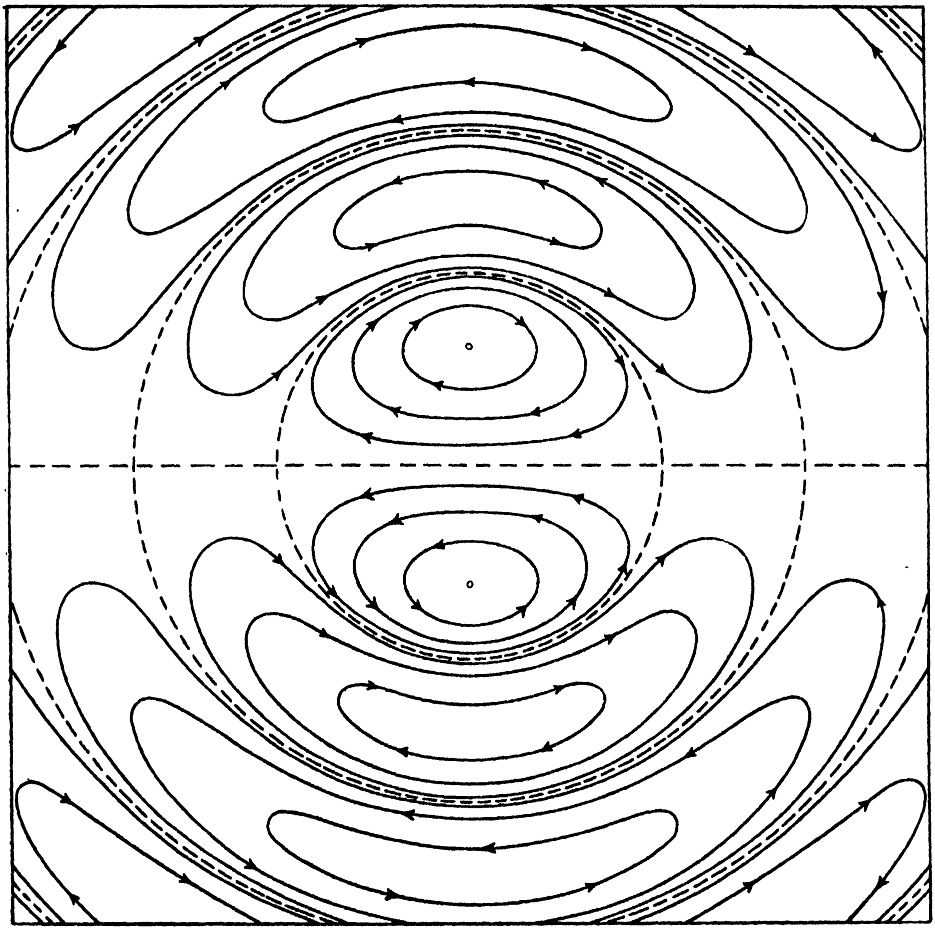


FIG. 4.—Stream-lines relative to the main body of the fluid in the disturbance due to a non-rotating core propagated along the axis of a rotating fluid.

and the equation of continuity as

$$\frac{\partial u}{\partial x}+\frac{\partial v}{\partial y}+\frac{\partial w}{\partial z}=0,$$

it will be seen at once that these equations are satisfied by

$$(u,\ v,\ w,\ p/\rho-\Omega^2\varpi^2)=(\mathrm{A},\ \mathrm{B},\ \mathrm{C},\ \mathrm{P})\,e^{i(ax+by+cz+nt)}$$

provided*

$$\left.\begin{aligned} ni\mathrm{A}-2\Omega\mathrm{B} &= -ia\mathrm{P} \\ ni\mathrm{B}+2\Omega\mathrm{A} &= -ib\mathrm{P} \\ ni\mathrm{C} \qquad\quad &= ic\mathrm{P} \\ \mathrm{A}a+\mathrm{B}b+\mathrm{C}c &= 0 \end{aligned}\right\}. \qquad (19)$$

---

* The terms involving products of u and v all vanish in this solution on account of the relation $\mathrm{A}a+\mathrm{B}b+\mathrm{C}c=0$, a condition which implies that the motion of the fluid is confined to the plane of the wave front.

Eliminating A, B, C, P it will be found that

$$a^2+b^2+c^2 = 4\Omega^2c^2/n^2. \tag{20}$$

Hence $$n/2\Omega = c(a^2+b^2+c^2)^{-\frac{1}{2}};$$

but $c(a^2+b^2+c^2)^{-\frac{1}{2}} = \cos\alpha$, hence

$$n/2\Omega = \cos\alpha. \tag{21}$$

The velocity of the waves represented by (18) is $n\lambda/2\pi$, where λ is the wave-length. Hence, from (21), the velocity is $\Omega\lambda\cos\alpha/\pi$, which is proportional to $\cos\alpha$ if λ is constant. It appears, therefore, that a spherical system of waves of length $U\pi/\Omega$ can be propagated with velocity U parallel to the axis of a rotating fluid.

It is an unusual feature of both the plane and the spherical types of wave that the amplitudes are not limited to small motions.

Experimental Demonstration of the Existence of a Non-rotating Sheath of Fluid Round a Sphere Moving in a Rotating Fluid.

It is difficult to realise a practical demonstration that any of the types of motion, considered above, actually exist in any real fluid; on the other hand, I have been able to show experimentally that, when a sphere moves along the axis of a rotating fluid, it is surrounded by a sheath of fluid which does not rotate with the rest of the fluid. That this result is to be expected is clear from equations (7), for at the surface of the sphere $f = 0$, so that $w = 0$, *i.e.*, the fluid immediately in contact with the sphere has no tendency to rotate it.

The apparatus with which this was demonstrated is shown in fig. 5. In this diagram G is a glass cylinder full of water, A is the axis about which the cylinder is made to rotate uniformly, B is the sphere which consists of a light celluloid ball, of the type used in the game of ping-pong. T is a thread which holds the ball down. The ball can be moved at a uniform speed along the axis by unrolling the thread, T, at a uniform speed from a reel R. This reel is mounted on a spindle fixed in the middle of a lid, C, which is fitted over the glass cylinder G. The thread passed from R through a series of small eyes, E, the last of which was in the centre of the bottom of the cylinder, and the ball, B, was painted with black and white quadrants, so as to make it easy to see whether it was rotating.

The apparatus was first rotated uniformly for some time till the ball and liquid were both rotating as a solid body with the glass cylinder. To perform the experiment, the reel, R, was suddenly fixed so that the thread, T, wound round the reel at a uniform rate. This gave the ball, B, a uniform speed along the axis of the cylinder.

It was found that the ball stopped rotating directly it started moving along the axis. As soon as the reel was released, so that the ball stopped moving along the axis, it quickly picked up the rotation of the rest of the system once more. To perform the experiment successfully, it was found necessary to use the lightest possible type of ball, to use a thin single or plaited silk thread, and to take great care that it was not twisted at the time of the experiment. To ensure success, it was found necessary also to make the ball move at a rate greater than about one diameter per revolution of the system (*i.e.*, $\mu < 2\pi$). If the ball travelled more slowly than this it was found that it did not stop rotating, and investigation of the stream-lines with coloured water, showed that a column of liquid, of the same diameter as the sphere, was apparently pushed along in front of the sphere. This observation suggests that the explanation of the rotation of the sphere, when it moves slowly along the axis, is that the stream-line, $\psi = 0$, does not keep to the surface of the sphere.

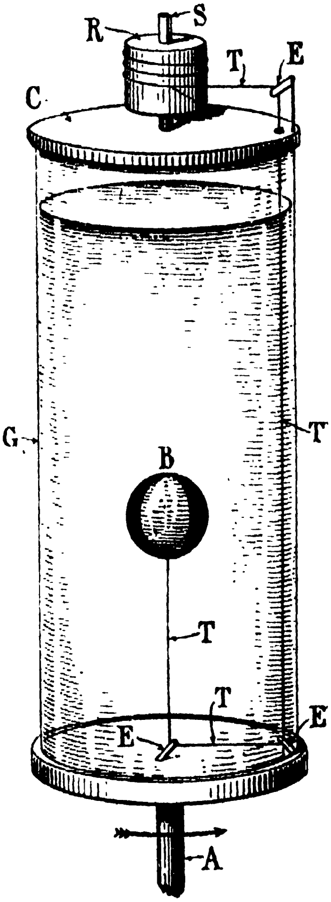


Fig. 5.—Apparatus intended to demonstrate that a sphere is surrounded by a non-rotating sheath of liquid when it travels along the axis of a rotating fluid.

In the course of these experiments, it was noticed that if the sphere was stopped suddenly when half-way up the cylinder, and if there was some colouring matter present to show up the motion, a mass of liquid appeared to detach itself from the sphere, and to continue moving along the axis of rotation with the same velocity as that with which the sphere had been moving. The impression produced on the author's mind was that the flow was similar to that shown in fig. 3. This conclusion, however, should be treated with reserve.

Polarisation of the Light Scattered by Mercury Vapour near the Resonance Periodicity.

By LORD RAYLEIGH, F.R.S.

(Received September 28, 1922.)

§ 1. *Introduction.*

When light belonging to the visual spectrum is scattered by air, as, for example, in the blue sky, there is no approximation to resonance between any period of vibration contained in the source and a free period of the scattering molecules or atoms. In such cases, there is an approach to complete polarisation of the scattered light, though a closer examination shows that the polarisation is in general incomplete. In argon, however, and probably in the other monatomic gases, there is a very near approach to completeness.*

In contrast with these cases is the scattering of light by mercury vapour in resonance with the source, when the latter belongs to the ultra-violet mercury line λ 2536, as emitted by a cooled mercury arc. This case was investigated by R. W. Wood.† He found that so long as the lamp ran under conditions favourable to producing an extremely narrow line without reversal, there was a copious emission of scattered light. I was anxious, 3 years ago, to trace the transition between this case and the ordinary scattering well away from the resonance periodicity. This was not fully achieved when my attention was turned to other matters, and shortly afterwards my laboratory at the Imperial College was dismantled. I have not since found an opportunity to resume, and it seems desirable to record the results, such as they are, without further delay.

§ 2. *Qualitative Detection of Polarisation.*

The mercury arc used was contained in a long vertical silica tube. A tungsten rod anode passed in at the top through a rubber connection and a side tube made connection with an exhausted glass flask, to afford a reservoir of vacuum, if the expression may be allowed. The cathode was the top surface of a barometric column standing in the tube, and the arc was struck by raising the reservoir till the mercury came in contact with the anode. A brass jacket with a quartz window was provided to keep the arc cool by water circulation, and the transverse field of a small electro-magnet kept the arc pressed against the front wall of the tube, so as to squeeze out the cooler layer of mercury vapour which otherwise intervenes.* The quartz tube was 8 mm. in diameter, and the current found to give the best result was about 10 ampères.

---

* 'Roy. Soc. Proc.,' A, vol. 98, p. 57 (1920).

† 'Phil. Mag.,' vol. 23, p. 689 (1912).

The light from this lamp was limited by a diaphragm about 1 cm. square, and made parallel by a quartz lens of short focus. The beam traversed a vessel exhausted of air and containing mercury vapour at the ordinary temperature. This vessel was made of square brass tubing, blackened inside, and had plane windows of silica glass, ground and polished. This material has a trace of double refraction,† but so very slight as not perceptibly to complicate the present experiments. The vessel was blackened internally, and the beam passed across the mouth of a cave or tunnel of some 10 cm. depth, which afforded a black background, as in previous experiments on scattering.

The first experiments were made with a view to detecting the existence of polarisation which Wood's early experiments had not revealed. It was thought that this might be attained if the light were filtered through a stratum of cold mercury vapour, in order to absorb the central constituent of the exciting spectrum line.

The polarising prisms used were a pair of quartz Wollaston prisms, made for me by Messrs. Hilger. For ultra-violet light it is necessary to avoid the use of Canada balsam, and the two halves of the prisms were put together in optical contact, like the echelon gratings made by the same firm. Only a very faint reflection could be seen from the interface. The two prisms were used in series, as the separation given by one alone was inconveniently small.

A camera with a simple quartz lens was mounted so that the axis of the latter was perpendicular to the primary beam traversing the vapour, and in front were the two Wollaston prisms set so as to separate the two polarised components in the up and down direction. In this way a photograph of the two components in the transverse direction could be taken, and it was found that the difference in intensity between them was not very conspicuous, in general agreement with Wood's early result.

To make a more searching test for polarisation it is usual to employ a Savart polariscope, *i.e.* a pair of plane quartz plates cut 45° from the axis and crossed. The fringes observed with this depend upon obliquity. Such an arrangement was not at hand, and I used a slightly wedge-shaped quartz plate, placed immediately in front of the window of the vessel, and therefore approximately in focus along with the primary beam as seen transversely. The principal directions of vibration in the quartz wedge were at 45° to the direction of the primary beam from the arc. Thus, as we go along the beam, we photograph it transversely through thicknesses of quartz giving first an odd and then an even number of half-waves' retardation; and if there is any polarisation of the scattered light we ought to get dark and bright bands alternating in the photograph of the beam. This was found to be the case, and the bands could be very distinctly photographed right up to the place where the beam entered the vapour. It is clear, therefore, that the light scattered by the resonant vapour is, after all, slightly polarised. This, however, does not go far to bridge over the contrast with the usual cases of scattered white light, which is almost completely polarised.

---

* Kerchbaum, 'Electrician,' vol. 72, p. 1074 (1914).

† 'Roy. Soc. Proc.,' A, vol. 98, p. 284 (1920).

§ 3. *Quantitative Tests. Filtration of Radiation by Mercury Vapour.*

To determine the amount of polarisation, I used a method which relies on the constancy of the mercury vapour lamp. It consists in first taking a photograph through the battery of double image prisms. The image in which the vibrations are vertical is the strongest. We now shift the plate so as to bring an adjacent part of it into use, and take another photograph with the lens stopped down, so that with the same exposure the more intense image has the same intensity that the less intense one had before. The diaphragm must be charged and successive photographs taken until this result is attained, as indicated by the photoelectric cell and galvanometer.* The ratio of intensities in the two images is then given by the inverse ratio of lens areas. The method assumes that the double image prisms do not favour one image more than the other. A selective absorption for ultra-violet light, analogous to that shown by tourmaline for visual light, would violate this condition. But it is known that quartz does not show any appreciable effect of this kind for the wave-length in question.†

As already remarked, the quality of the radiation will vary as the beam traverses the column of mercury vapour. The kind of radiation which is most copiously scattered belongs, as Wood has shown, to a very narrow spectral range, and is rapidly removed from the primary beam. Further along, the scattering is much diminished in intensity.

By changing the thickness of mercury vapour traversed, we may, to some extent, examine how the polarisation of the scattered light varies as the resonance frequency is departed from. It is probably a matter of indifference whether the primary or the scattered beam traverses the absorbing column. In my experiments the absorption was chiefly exerted on the primary beam.

---

* 'Roy. Soc. Proc.,' A, vol. 97, p. 440 (1920).

† With calcite prisms it is appreciable, and gave much trouble until detected.

It is impracticable to allow the primary beam to pass very close to the observing window; if it did so, there would certainly be false light from the latter, so that a short intervening column of mercury vapour cannot be avoided. One of the vessels used was arranged so that the entrance window for the primary beam and the observation window at right angles to it came together at an edge where they were made tight with a little cement. The primary beam was limited by a slit parallel to this edge, and the scattering could be examined immediately it entered the vapour, with only about 4 mm. thickness of mercury vapour for the scattered radiation to traverse.

In this case the weaker component polarisation was measured as 90 per cent. of the intensity of the stronger one.

Another determination was made, using a vessel in which 20 cm. of mercury vapour were traversed by the primary beam, and 7·5 cm. by the scattered beam before coming to the window. In this case the mean of several fairly accordant determinations gave 60 per cent. intensity ratio as compared with the 90 per cent. of the previous case.

The exposure found desirable was sixty times that used before, which will serve to give a rough idea of how greatly the intensity of scattered radiation is reduced by filtering the beam through 27 cm. of mercury vapour. It will be noticed that the mercury vapour used was *in vacuo*. Mercury vapour in presence of air gives a broader absorption band. It was not used in these experiments.

To get some idea of the spectral range which is approximately cut out by 27·5 cm. of mercury vapour, the structure of the line 2536 from the lamp was examined by a quartz Lummer plate. The fringes were projected by a quartz-fluorite achromat on to the slit of a small quartz spectrograph, in order to separate this line from the others. It appeared that the effect of interposing a column 27·5 cm. long of mercury vapour at room temperature was to cut out the most intense portion of the line having a breadth of 1/100 of an Ångstrom, which was about 1/5 of the entire breadth of the line. The structure of the line appeared to be unsymmetrical. The photographs were definite enough, but more experience than I can claim would be required to interpret them with full confidence.

§ 4. *White Light Scattered by Dense Mercury Vapour.*

It seemed desirable to make certain that mercury vapour in scattering white light would behave like air and the ordinary gases. For this purpose it was necessary to use dense vapour, since the scattering is so much less powerful than in the case of ultra-violet resonance. The vapour of mercury boiling or nearly boiling at atmospheric pressure was used. The vessel was one of cross shape, made of steel tube acetylene welded, the same, in fact, as that used for my experiments on scattering by carbon dioxide at high pressure.*

The light from a carbon arc was made parallel by a lens and entered by a mica window, and the scattered light was observed through a window of silica glass, ground and polished. These windows had the advantage of not cracking, which glass windows were very apt to do under the conditions used. The windows were merely held against the ends of the steel tubes, without packing or cement. The whole was placed in an electrically heated oven, with holes corresponding to the windows, and special heaters were arranged to make sure that the windows were hotter than the rest of the vessel, so that no mercury could condense upon them. The vessel was open to the air through an exit tube which served as a reflux condenser.

A little mercury was poured in over-night, and the heating current turned on. By morning, all the liquid mercury had apparently gone, but on turning on the arc it was observed that the track of the beam was of sky-blue colour† and much more intense than in dust-free air. At the same time the polarisation was approximately complete. So far as visual observation could show, all the light was concentrated in one of the images formed by a double image prism.

For whatever reason, dust was not observable in this experiment. The high temperature seems in some way to clear it off.

On adding more liquid mercury, ebullition began, and the disturbance caused dust to rise with the vapour. The individual dust particles could be seen, comparatively yellow in colour, showing by contrast against the blue background formed by the gaseous atmosphere of mercury. The yellow particles showed in both images, but the blue background only in one, as before. Reflex condensation could be seen in the glass exit tube all the time.

The heating current was switched off, and the dust cloud slowly diminished, leaving a blue beam as before. It was confirmed by photography that the light was polarised with approximate completeness. The experiment was not pushed to the point of examining the small defect of polarisation if any. It was considered to establish definitely that mercury vapour behaves to a first approximation like ordinary gases in scattering polarised white light at right angles to the incident beam.

§ 5. *Possible Explanation by Rotation of Luminous Centres.*

Very little is known about the rotation of molecules, and according to the classical mechanics the monatomic mercury molecule, which has a ratio of specific heats 1·66, should have no rotational energy. No one, however, would be disposed to press such a conclusion without reserve at the present time, and it may be noticed that a rotation of the emitting centre about the line of vision would destroy the polarisation of the emitted light, if we suppose the emission to be maintained during a rotation of 90° or more, and the direction of vibration to be fixed within the molecule. In the absence of any notion of the angular velocity, we cannot say what duration would be required, but it seems conceivable that in the case of exact resonance the duration might be long enough for this cause to be effective. It was attempted to bring it into evidence by causing the mercury vapour to move as rapidly as possible across the direction of the primary beam, but without effect. There was no tendency for the source of secondary emission as a whole to move downstream.

---

* 'Roy. Soc. Proc.,' A, vol. 95, p. 173 (1918).

† At least to my eyes.

In these experiments the vessel was a vertical brass tube about 2 cm. square in section, with entrance and observation windows in adjacent sides, and at the same level. These windows were of silica, cemented on with sealing wax, which was found to stand the moderate temperatures used in distillation. Above the level of the windows was an annular gallery containing mercury, the walls of the tube forming the inner railing of the gallery. Below, the closed end of the tube contained granular charcoal, and was immersed in liquid air. This formed the condenser. The tube was thoroughly exhausted with an air pump, and the charcoal kept the exhaustion good, in spite of any tendency of the blackened walls to give off gas. These various conditions were designed to give as free a distillation as possible of rare mercury vapour down the tube.

Photographs were taken (1) without cooling the charcoal (no distillation); (2) charcoal cooled, mercury at room temperature; (3) charcoal cooled, mercury as hot as the hand could bear.

In all these cases the beam was photographed stretching straight across the vessel, and sometimes when the exposure was long a diffuse luminosity surrounded it, due, no doubt, to tertiary radiation. But in no case was any dissymmetry observed, such as would be produced if a luminous source moved downstream an appreciable distance during emission. It may be assumed that the mercury vapour was moving with molecular velocity, or, at least, with a velocity of that order.

Experiments on somewhat similar lines were made by Phillips,* and he obtained a visual green glow and some ultra-violet radiation of λ 2536 with it. This could be made to go a long distance. The effect, however, appears to be a secondary one, since it is only a small part of the total radiation λ 2536 which can have its origin displaced. To bring out Phillips' effect requires a more vigorous distillation than could be set up in the metal vessel above-mentioned with flat windows and sealing-wax joints. The advantage of this construction is in avoidance of stray light. Phillips used a silica tube strongly heated, so as to give abundant distillation. Such a tube will stand the heat, but is much less favourable in respect of stray light.

---

* 'Roy. Soc. Proc.,' A, vol. 89, p. 39 (1913).

It looks as if the centres capable of being excited to durable luminosity were few, and that much mercury must be passed in order to obtain a sufficient supply of them. But there is still a great deal to be done in elucidating the effect, and I do not wish the above suggestion to be taken as more than tentative.

§ 6. *Conclusion.*

Finally, it is to be remarked that the resonance radiation of mercury is completely lost if the vessel containing the vapour is not exhausted of air. There is no reason to suspect that anything analogous to this occurs when a mixture of ordinary gasses is scattered by white light, and certainly existing theories of the latter phenomenon do not contemplate anything of the kind. As long as this peculiarity remains unexplained, some reserve is necessary in treating resonance radiation as a case of ordinary scattering.

The chief results of the present investigation are :—

(1) White light scattered at right angles by dense mercury vapour is, to a first approximation, completely polarised.

(2) Ultra-violet radiation of the mercury spectrum line λ 2536, when examined immediately it enters mercury vapour in an exhausted vessel at room temperature, gives a scattered radiation which is slightly, though definitely, polarised.

(3) This polarisation has been observed to increase as the beam is filtered by penetration of a considerable depth of vapour. After penetration of 27·5 cm. of vapour the weaker polarised image had 60 per cent. only of the intensity of the stronger one, instead of 90 per cent. as at first. The radiation removed by the filtration appears to lie within a spectral range of about 1/100 Ångstrom.

The Scattering of Hydrogen Positive Rays, and the Existence of a Powerful Field of Force in the Hydrogen Molecule.

By G. P. Thomson, M.A., Fellow of Corpus Christi College, Cambridge,
Professor of Natural Philosophy in the University of Aberdeen.

(Communicated by Sir J. J. Thomson, F.R.S. Received August 5, 1922.)

1. *General Scheme of Research.*

In view of the extremely important results obtained by Sir E. Rutherford and others from a study of the scattering of α-rays, it seemed worth while to investigate the scattering of particles moving with smaller velocities such as occur in the positive rays. The most interesting, because the simplest, are the rays of positively charged hydrogen atoms, which presumably consist simply of a nuclear particle, or proton. The experiments described in this paper were made in some cases with these rays, in others with the positively charged hydrogen molecules, systems consisting of two protons and one electron. The scattering medium was in all cases hydrogen gas. This was chosen largely for convenience, as the experimental arrangement is considerably simplified if the same gas is used to produce the rays and to scatter them, and also because, with the exception of helium, the molecule of hydrogen is the simplest known, and there seemed more hope of obtaining results which could be given a definite theoretical interpretation.

The general scheme of experiment was to produce the rays in a discharge tube, analyse them by magnetic and electric fields in the ordinary way, cut off all except those of the kind required by a slotted diaphragm, pass the remainder through a chamber containing the scattering gas, and receive them in a Faraday cylinder arranged behind a slit of variable width. The experiment consisted in finding how the charge received by the Faraday cylinder varied with the width of the slit, when this was made wider than the geometrical "shadow" of the slot in the diaphragm. Any rays lying outside this "shadow" must have been scattered.

The energy of the rays was of the order of 10,000 volts; the pressure of gas in the scattering chamber varied from 1·5 to 12 thousandths of a millimetre of mercury. The scattering observed was of the order of 1°.

The outstanding result of the experiments is that the scattering is many times (10—20) greater than would be expected, on the assumption that the only forces acting are the ordinary forces of electrostatics varying as the inverse square.

2. *Experimental Arrangement.*

The apparatus used is shown in general arrangement in fig. 1, and the slit mechanism in more detail in fig. 2.

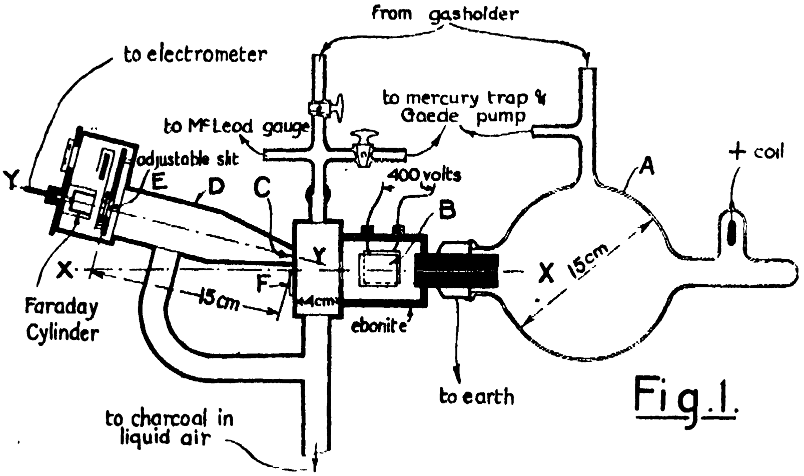


Fig. 1.

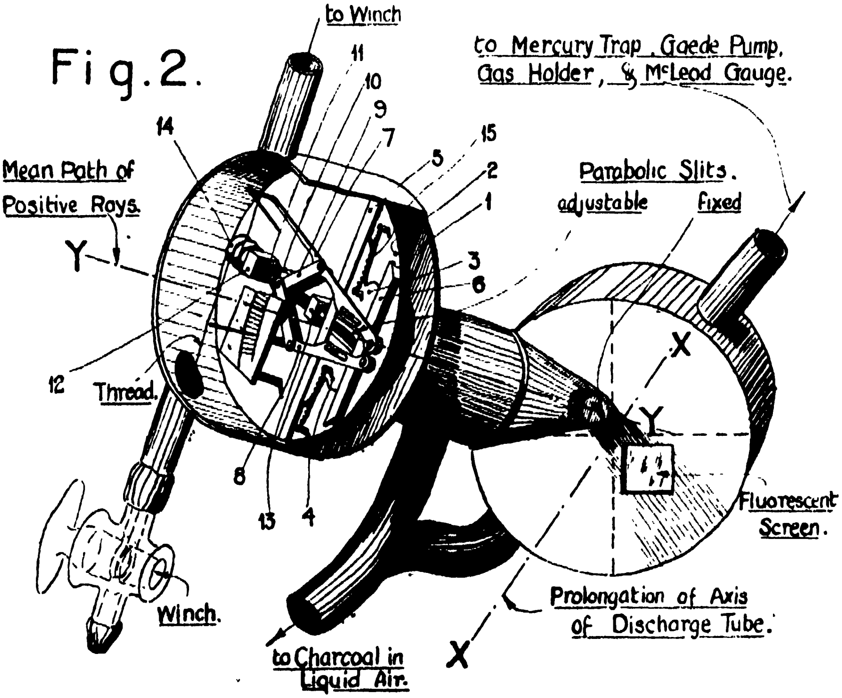

to Winch
11
10
9
7
to Mercury Trap, Gaede Pump, Gas Holder, & McLeod Gauge.
14
15
5
2
1
Mean Path of Positive Rays.
Parabolic Slits.
adjustable
fixed
Y
3
6
X
12
Thread.
Y
8
13
4
Fluorescent Screen.
Winch.
Prolongation of Axis of Discharge Tube.
to Charcoal in Liquid Air.
X

Fig. 2.

In fig. 1 A is a discharge bulb connected to a Gaede pump as shown, and fitted with a concave cathode pierced by a fine tube of 0·37 mm. diameter and 5·1 cm. length.

The rays passing through this fine tube are analysed by electric and magnetic fields between the plates, B. They then fall on the plate, C, pierced with a parabolic slit 1·2 mm. in width. By adjusting the magnetic field any required type of rays (in practice hydrogen atom or molecule) can be made to pass through the slit and into the scattering chamber, D. The pressure here is kept low by the use of charcoal cooled by liquid air, which also has the effect of removing impurities from the hydrogen, or slight traces of vapour given off from the apparatus. The rays then fall on the adjustable slit, E, which is arranged parallel to the parabolic slit in C. After passing E they are collected in the Faraday cylinder. The width of the slit, E, can be adjusted without disturbing the vacuum by means of two winches, fig. 2, of the type invented by Dr. Aston for lowering a plate in the ordinary photographic type of positive ray apparatus.

The slit mechanism, shown in detail in fig. 2, was designed and made by Mr. W. S. Farren, M.A., and has worked with complete success. In the figure 1 is the airtight face of the chamber, 2 is a base plate, 3 and 4 are the plates which form the slit and move in guides, 5 and 6. They are actuated by slotted levers, 7, 8, moved by links, 9, which are connected to the nut, 10. This is driven by the screw, 11, which is turned by two threads round the pulley, 14, and kept from axial motion by the stops, 12 and 13. All backlash is taken up by the springs, 15, which are always in tension. The width of the slit is determined by reading through a plate-glass window the position of two pointers on the scale as shown, the scale being fastened to a mirror to obviate parallax.

The procedure in an experiment was as follows:—The apparatus was pumped out and a charge of hydrogen admitted from the gas holder. It was then pumped down again till the pressure was a few hundredths of a millimetre, and the taps turned which cut the scattering chamber off from the discharge. The pressure in the scattering chamber was then reduced by means of the charcoal, while the discharge tube was pumped down until a balance was obtained between the pump and a leak through a fine capillary to the gas holder, which was adjusted to give a pressure of the order of 1/100 mm. in the discharge tube.

When the discharge had been running for some little time and the pressure seemed steady, the spots were examined on the willemite screen, F, on to which they were thrown by reversing the field. If these were satisfactory in appearance, the rays were thrown on to the slot in C and so into the Faraday cylinder. This was connected with a Dolazelek electrometer (period 50 seconds), and the current obtained from the rays balanced against a high resistance of xylol and alcohol (resistance of order 10^{12} ohms). The deflection of the electrometer was read on a scale in the ordinary way. A deflection of about 700 mm. corresponded to 1 volt. The slit, E, fig. 1, was first adjusted to a width of about 1·5 mm., and the magnetic field adjusted till the electrometer deflection was a maximum. This ensured that the pencil of rays was correctly centred on the slot, C. The width of E was then gradually increased and the readings of the electrometer taken for various widths of E. At the same time the pressure of the scattering gas was read on a McLeod gauge. The slit was gradually opened and closed several times in the course of an experiment, which might last about three-quarters of an hour. Any irregularity in the running of the coil would, of course, lead to inconsistent results being obtained, and probably slight irregularities of this kind constitute the most important of the experimental errors. It was found, however, that so long as the Cox brake gave a uniform sound, the results so obtained repeated with considerable accuracy over the course of an experiment (see fig. 3). It was found advisable to place a fairly strong electromagnet near the Faraday cylinder. This considerably reduced the deflection, part of which is, no doubt, due to δ-rays caused by the impact of the positive rays. The magnet prevents these from escaping from the cylinder. It was always used during the experiments recorded, and was so strong that further increase of strength had no effect.

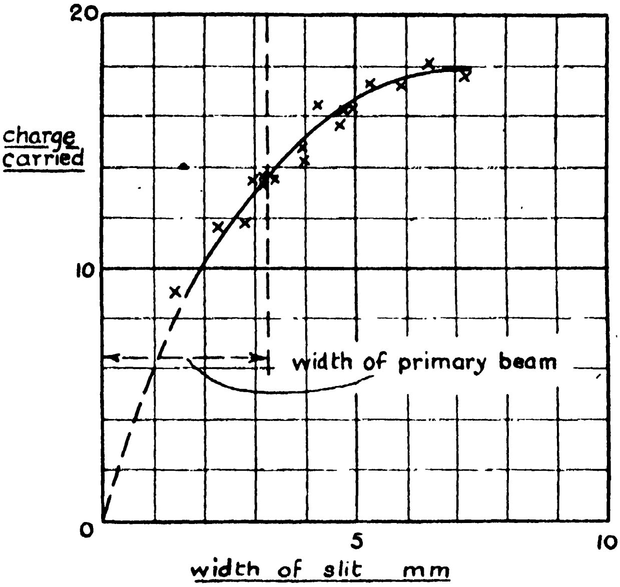
3. *Precautions.*

When it became clear that the scattering observed was greater than would be expected theoretically, a variety of precautions were taken to eliminate possible spurious scattering effects. In the earlier experiments, the action of the charcoal was relied on to purify the gas in the scattering chamber from mercury, of which the spectroscope showed a good deal in the discharge tube. In case this was not really adequate, a mercury trap dipped in liquid air was fitted between the pump and the discharge tube. In this way the mercury was removed until no trace of its spectrum could be seen in the discharge. It was found, however, that there was no perceptible difference in the scattering curves obtained before and after the use of the mercury trap, which was, however, continued as an extra precaution. It was suggested that, in spite of the charcoal, the vapour pressure of the tap grease used might be sufficient to produce an appreciable effect. This possibility is, however, excluded by the fact that perfectly sharp photographs are regularly obtained from apparatus in which winches, lubricated with the same tap grease, are used. Great care was taken to clean the slot at C, but it seemed possible that some minute quantity of dirt might become charged, and by its repulsion of the rays produce an appearance of scattering. To test this, at the end of the research the slot was greatly enlarged so that there was no obstruction anywhere near the rays, but no great difference was found as a result, the only change being a slight *increase* in the scattering over big angles, due, no doubt, to the fact that rays of other kinds were no longer excluded from the scattering vessels.

4. *Readings.*

At the pressures at which the discharge runs best the molecular rays are considerably more numerous than the atomic, and so the results obtained from them are probably more accurate than those obtained from the atomic ones. However, quite sufficient results were obtained with the atomic rays to show that the scattering was nearly the same in the two cases, indeed, it is probably identical within the limits of accuracy of the present apparatus.

A typical curve, with the experimental points marked, is shown in fig. 3.



Molecule. Mean energy 11000 volts
Pressure 4·5 $\frac{\text{mm}}{1000}$
Fig. 3. Mercury trap used.

In every case the slit was opened and closed several times during the experiment, so that any effect due to a gradual change in the discharge conditions is excluded. The curves show deflection of the electrometer plotted against *total* width of the slit; in order to find the relative number of particles falling on different portions of the slit, it is therefore necessary to find the slope of the curve at the various points. The dotted ordinate shows the width of the slit which could be reached by unscattered particles, and the rise of the curve to the right of this ordinate represents scattering. Owing, however, to the circular section of the fine tube, the distribution of the rays would be by no means uniform apart from scattering. In the area bounded by the dotted ordinate, probably the great bulk of the unscattered rays would fall within considerably narrower limits.

Reduction of Results.—The ideal in an experiment of this kind is to deduce from the quantities measured the probability at each encounter of any specified angle of scattering. Though the data obtained are theoretically sufficient for this to be done, the results with the present experimental arrangement would be so uncertain that no useful purpose is served by performing the reductions. It is hoped, however, by certain modifications of the apparatus, to reach a degree of accuracy which will make this possible in future experiments. The present work must be taken as determining the general magnitude of the effect to an accuracy of 10–20 per cent., and the interest lies in the very large discrepancy between the value obtained and that which would be expected on the ordinary theory (*vide infra*).

In order to compare the results of experiments made at different pressures, it is desirable to find a single quantity which can be regarded as measuring the scattering. If the curves are of different shapes, as one would theoretically expect them to be, it is difficult to find any one quantity which can fairly be compared. Perhaps the best is to take the fraction scattered through a distance greater than some arbitrarily chosen value, but care must be taken in selecting this value. Taking a value which roughly corresponds to the width of the main portion of the unscattered beam, the results obtained are shown in fig. 4. The points lie on a curve which is fairly parabolic for large pressures, but more nearly linear for small pressures. This might be expected. If the number of collisions which is effective in producing measurable scattering rarely exceeds one, the scattering will be directly proportional to the pressure. If the number of such collisions is large, the scattering on the usual probability theory will be proportional to

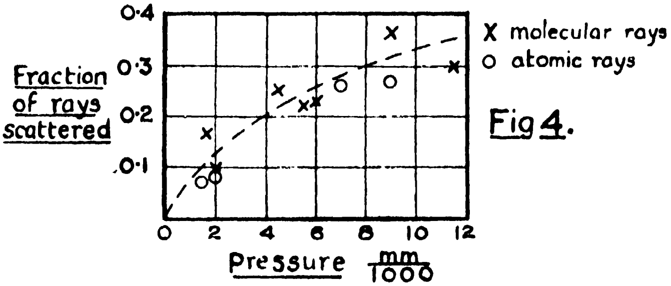


Fig 4.

the square root of the number, and so to the square root of the pressure, giving a parabola. It must not, of course, be assumed that every collision, as calculated from kinetic theory data, will give appreciable scattering, but this number will give an upper limit. However, if this explanation of the shape of the curve is correct, it is perhaps worth noticing that the curvature begins to be marked somewhere near the point where the average particle will have made more than one kinetic theory collision (allowance being made for the speed and size of the rays).

In order to get a more definite measure of the scattering, it was decided to compare the mean curve from four experiments, at not very different pressures, with the curve obtained at the lowest pressure available, modified slightly by exterpolation, to represent as nearly as possible the unscattered beam.

The two curves are shown drawn out in fig. 5, the scales being adjusted to make the total within a slit width of 0·75 mm. come the same. The next stage in reduction would be to take the slopes at various points. It would then be necessary to find from the actual scattering curve the ideal curve due to an infinitely narrow beam. Though this is theoretically possible, it was found to involve such a loss of accuracy as to give a result that was of little value, so the work was not proceeded with, especially as it would then have been necessary to allow for the fact that the scattering occurred throughout the volume of the chamber instead of being concentrated in a

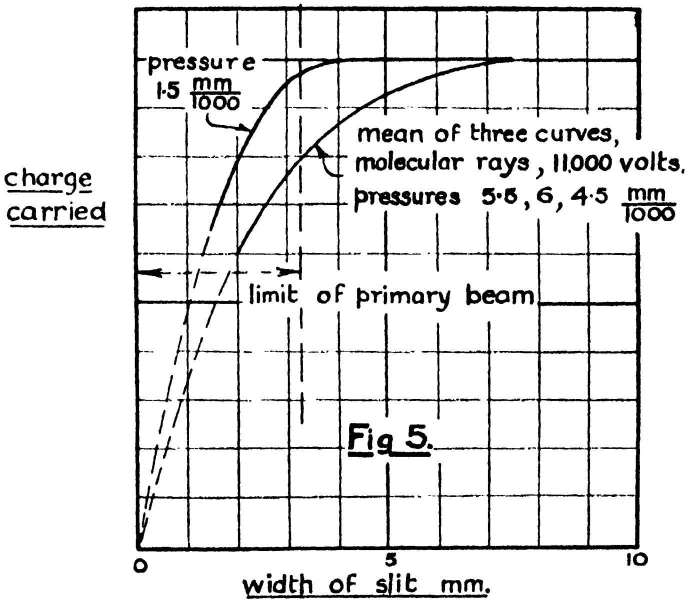


Fig 5.

thin layer, and was symmetrical in two dimensions instead of being in one as measured.

For these reasons, the most satisfactory way of comparing theory with the experiments is to deduce what shape of curve would theoretically be expected, and compare it with those obtained experimentally. Consider first the case of a hydrogen nucleus moving through hydrogen gas. The rays we are dealing with have an energy of roughly 10,000 volts, and the smallest scattering we can detect is about 1/300 radian, the largest about 3/100. The possible scattering particles in the hydrogen molecule are the two nuclei and the two electrons. To a fair approximation the nuclei can be regarded as acting independently, the connection between them being ignored. This can be seen as follows: the sideways velocity given to the nucleus by the particle, if the former is free to move, is equal and opposite to that which the particle itself receives, since they are of equal mass. The time of a collision is d/v, where v is the velocity of the ray and d is the distance over which the forces of the collision are important, which we may take as of the order of the diameter of the molecule. The velocity given to the nucleus is θv, where θ is the angle of scattering in radians. Hence the distance which the nucleus would move, if unconstrained by the rest of the molecule, is approximately $\frac{1}{2}\theta v . d/v = \frac{1}{2}\theta d$. Since θ is small, it follows that the displacement of the nucleus which would occur if it were unrestrained is a small fraction of the molecular diameter. Originally the nucleus was in equilibrium, and the force due to this small displacement from equilibrium will not be comparable with that between the ray and the nucleus (which approach within a distance less than $d/2$), unless the law of force between the constituent parts of the molecule is one which varies very rapidly with distance. It seems reasonable, therefore, to treat the nucleus as free to move.

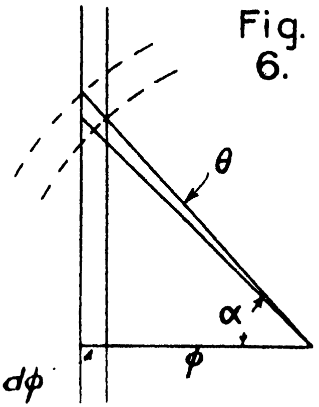


Fig. 6.

If a particle of mass m moving with velocity v shoots past an equal particle, which it repels with a force ϵ^2/r^2, the angle through which the direction of motion of the moving particle is deviated, supposed small, can be shown to be $2\epsilon^2/mv^2p$, where p is the perpendicular distance from the second particle on the original line of motion of the first. In our case this may be written ϵ/Vp, where V is the potential difference through which the charge ϵ has fallen.

Now, in the experiments, the scattering is measured through a slit, while θ

is the radial scattering. Consider a screen at unit distance from the scattering particle (fig. 6). The chance of a ray being scattered between θ and $\theta+d\theta$ is the chance of its coming between p and $p+dp$ of a nucleus, and is $2\pi p\mathrm{N}dp$ per unit length, where N is the number of nuclei per unit volume. But $2\pi p\mathrm{N}dp = 2\pi \frac{\mathrm{N}\epsilon^2}{\mathrm{V}^2}\frac{d\theta}{\theta^3}$, since $p = \frac{\epsilon}{\mathrm{V}\theta}$; thus the surface density on the screen of particles scattered through an angle of θ is $\frac{\mathrm{N}\epsilon^2}{\mathrm{V}^2\theta^4}$. Hence the chance of a ray being scattered into the strip between ϕ and $\phi+d\phi$ is $2d\phi\int_0^{\pi/2}\frac{\mathrm{N}\epsilon^3}{\mathrm{V}^2\phi^4\sec^4\alpha}\,.\,\phi\sec^2\alpha\,dd = \frac{\pi}{2}\frac{\mathrm{N}\epsilon^3}{\mathrm{V}^2\phi^3}\,d\phi$, where $\phi = \theta\cos\alpha$.

In the actual experiment the scattering occurs at all points along the path of the rays, and the linear displacement of a ray at the slit will depend both on the angular deviation caused by the collision and on the distance from the slit at which the collision occurs. Without some limitation on the law we have been considering, it would be necessary to consider a large number of very small collisions. But these small collisions correspond to cases in which the ray is always at a considerable distance from the nucleus. If the ray is sufficiently distant, the forces will be small enough for the molecule to act almost as a rigid body, and, as it is neutral, the resultant force will be small. For these reasons it is desirable to limit the "collisions" which we consider to those in which p is less than a certain arbitrary quantity, which we will take as the radius of the hydrogen molecule, or about $1{\cdot}2\times10^{-8}$. We shall see later that the precise value taken does not affect the main conclusions. Let us call this value r, and the chance of a collision in a length dx, λdx, thus, $\lambda = \mathrm{N}\pi r^2$. The fraction of the total number of collisions which results in a scattering between ϕ and $\phi+d\phi$ is $\epsilon^2\,.\,d\phi/2r^2v^2\phi^3$ as long as $\phi > \epsilon/v_r$. (If ϕ is less than this value, θ, which is $\phi\sec\alpha$, may be less than ϵ/v_r, and the expression will be complicated, but since scattering as small as this could barely be detected, its effect will be ignored.) Then if l is the length of the scattering chamber (15 cm.), the fraction which gets through without any collision is $e^{-\lambda l}$. The fraction which make one collision only is $\lambda le^{-\lambda l}$. Now, at the pressures actually used, comparatively few will make more than one effective collision. (It should be noticed that three-fourths of the collisions for which $p<r$ will give a value of $\phi<\epsilon/v_r$ after resolving perpendicular to the slit.) We can approximate to the effect of the few multiple collisions by regarding them simply as so many extra single collisions. Now the total number of collisions is λlx (the number of rays), so we shall suppose that λdx collisions occur in the interval dx, each of which is to be treated as a single collision.

The fraction $\frac{\epsilon^2 d\phi}{2r^2V^2\phi^3}$ of these make an angle between ϕ and $\phi+d\phi$ after collision. If y is the linear deviation at the slit and x is measured from the beginning of the scattering chamber $y=(l-x)\phi$. Hence a fraction of the rays equal to $\frac{\lambda\epsilon^2\,d\phi\,dx}{2r^2V^2\phi^3}$ lie within a width y and $y+dy$ where $dy=\phi dx$ and ϕ has all possible values. Thus the total fraction between y and $y+dy$ is $dy\int_{y/l}^{+}\frac{\lambda\epsilon^2 d\phi}{2r^2V^2\phi^4}=\frac{\lambda\epsilon^2 l^3}{6r^2V^2y^3}dy$, the upper limit of integration being an angle so large that the amount scattered through it can be neglected. Here $y/l>\epsilon/Vr$. Substituting for λ we get $\frac{\pi\epsilon^2 Nl^3}{6V^2y^3}$.

To allow for the finite width of the beam, we will suppose it of width $2a$, the intensity being uniformly distributed. If $2a$ is taken as the actual width over all, this will certainly give an over-estimate of the spreading to be expected, as in practice the intensity is much greater at the centre.

Let the strength of the beam be I, then the amount in a range of width $d\eta$ is $\mathrm{I}d\eta/2a$. The intensity at a point distant y from the axis is

$$\int_{-a}^{a}\frac{\pi\epsilon^2Nl^3}{6V^3(y-\eta)^3}\cdot\frac{\mathrm{I}}{2a}\,d\eta \text{ if } y>a+\frac{l\epsilon}{Vr}$$

and

$$\int_{-a}^{y-l\epsilon/Vr}\frac{\pi\epsilon^2Nl^3}{6V^2(y-\eta)^3}\cdot\frac{\mathrm{I}}{2a}\cdot d\eta+\frac{\mathrm{I}}{12a}\left(a-y+\frac{l\epsilon}{Vr}\right)\frac{\pi VNr^3}{\epsilon} \text{ if } a<y<a+\frac{l\epsilon}{Vr}.^*$$

By simple integration the amount outside a can be shown to be $\frac{\mathrm{I}\pi N\epsilon^2l^3}{12V^2a}\left\{\frac{3}{2}\frac{Vr}{l\epsilon}-\frac{1}{4a}\right\}$ provided $2a>\frac{l\epsilon}{Vr}$ as was the case in these experiments. The amount outside $a+\xi$ where $\xi>\frac{l\epsilon}{Vr}$ is $\frac{\mathrm{I}\pi N\epsilon^2l^3}{12V^2\xi(2a+\xi)\xi}$ and so depends only indirectly on r.

Taking the value of N corresponding to a pressure of 5/1000 mm., and treating the two nuclei in the molecule as independent, these expressions become $\mathrm{I}\times1{\cdot}5\times10^{-2}$ if $a=0{\cdot}17$ cm. and $\mathrm{I}\times3{\cdot}3\times10^{-3}$ if $a+\xi=0{\cdot}20$ cm. The value of $l\epsilon/V_r=0{\cdot}016$.

The corresponding values read off the curve of fig. 5 are $\mathrm{I}\times1{\cdot}9\times10^{-1}$ and $\mathrm{I}\times1{\cdot}3\times10^{-1}$. Thus the observed scattering exceeds that calculated on the assumption of the inverse square law by a factor of 13 in one case and 40 in the other.

The experiments in fig. 5, however, were made with molecular rays. As before, we may regard the nuclei in the positively charged molecule as acting independently. The effect of this is nearly to double N, but, on the other hand, after the collision the difference in velocities of the nuclei may result in the molecule being broken and the chance that the electron will go with either nucleus is probaby about even. Thus half the scattered particles will be neutral, and so will not be detected in the Faraday cylinder. The scattering of the remainder will be reduced by the force of the bond to an extent which it is difficult to determine; but if this change is small, the other two effects will roughly balance, and we should expect equal scattering for atomic and molecular rays, as indeed the experiments indicate. If the molecular ray is not broken, a given collision will only give it half the deflection which an atomic ray would have received. Thus we shall have twice the number of deflections, each of half the amount. With the law of force we are considering, it can be shown that this means that half as many are scattered per unit length through any specified angle. On this supposition, then, we should expect the scattering of molecular rays to be half that of atomic.

---

* Unless $2a<l\epsilon/Vr$ when the limits are $l\epsilon/Vr-a<y<a+l\epsilon/Vr$ and the intensity is $\mathrm{I}V\pi Nr^3/6\epsilon$ for $a<y<l\epsilon/Vr-a$, but this case does not occur in practice.

The investigation above has taken no account of the electrons as scattering agents. Owing to their small mass, their direct effect is negligible and they can only act if they are so tightly bound to the nuclei that they form a practically rigid body with them for forces of the order we are considering. This seems improbable, and even if it were the case it would only increase the atomic scattering by a factor of 2, which is quite insufficient to account for the discrepancy. The velocities of the electrons on the Bohr theory is only of the same order as that of the rays, so the chances of collision are not greatly increased on this account.

It is known that, as a result of collisions, the positively charged hydrogen atom and molecule can pick up an electron and be neutralised, while the atom can even acquire a negative charge. Such neutralised rays would, of course, not be detected in the Faraday cylinder. This will reduce the total effect measured by the electrometer, but it does not seem probable that it will much affect the value found for the scattering. A ray which has been scattered will have the same chance of being neutralised as one which has not, and, since the scattering must be almost entirely the result of action between the nuclei, a ray that has been neutralised will have the same chance of being scattered as one that has kept its charge. Thus the scattered and unscattered rays will be reduced by neutralisation in the same proportion. It will, of course, be understood that in the above experiment we are only concerned with neutralisation in the scattering chamber. Rays which were neutral when passing the electric and magnetic fields would not be deflected and would not pass into the scattering chamber at all.

In order to account for the discrepancy it is necessary to suppose that there exists a field of force in the hydrogen molecule which, at distances of the order of 10^{-8} cm. from a nucleus, is many times stronger than the ordinary electrostatic force, and is capable of acting on a positively charged particle. This is not in conflict with Rutherford's results on the scattering of α-rays, as, owing to the much greater energy of the α-rays, a field of force amply sufficient to account for the above experiments would have no appreciable effect in their case. In addition, he is concerned with the interaction of two particles, one of which, the α-ray, contains exactly twice as many protons as electrons, and the other, the nucleus of the scattering atom, approximately twice as many. It is possible that this special relation may produce an almost complete compensation.

The field of force required may be due to an actual change in the law of force between nuclei at distances of this order; or if this law is taken as always being the inverse square, it must be the result of the mechanism, whatever it is, which keeps the molecule in stable equilibrium.

Note on "Stossstrahlen."

Glimme and Koenigsberger,* as a result of measurements with Faraday cylinders and thermopiles, believe that they have proved that in a beam of hydrogen positive rays, at pressures of the same order as those used in the above researches, the number of rays as measured by the Faraday cylinder is much greater (3–20 times) than that as measured by the thermopile. They interpret this as meaning that the large majority of the rays detected by the Faraday cylinder consist of slowly moving rays, which they call "stossstrahlen," and believe to be produced by impact from secondary rays, which themselves are produced by impact of the original beam. They give the energy of these slow rays as less than 200 volts. If such were the case, the interpretation given above of my experimental results would, of course, have to be considerably modified. Their view seems hard to reconcile with the fact that a beam of rays, even at comparatively high pressures, shows as a definite pencil, with fairly sharp edges. To test it further, I made an experiment in which a beam of rays was passed between a pair of parallel plates before falling on a Faraday cylinder. The plates could either both be earthed, or one could be earthed and the other raised to a known potential, large enough for all rays with energy as little as 200 volts to be swept on to the plates, and so prevented from reaching the Faraday cylinder. At a low pressure (3/1000 mm.) this caused no appreciable difference in the charge received by the cylinder. At a higher pressure the charge could either be increased or decreased, depending on whether the sign of the high potential plate was such as to drive positive or negative ions into the Faraday cylinder. But even so, the diminution in the charge observed was in no case comparable with that which would have occurred if there had been anything like the proportion of slow rays which Glimme and Koenigsberger require to explain their experiments. From theoretical considerations such an effect seems very improbable. Thus even if a ray is deviated through an angle of 1 in 30 (and both theory and experiment show that very few receive a greater deflection) the scattering particle, if of equal mass, would only acquire 1/30 of the velocity and 1/900 of the energy of the original ray, and would move off almost at right angles to the beam of rays. With an energy of 10,000 volts this gives 11 volts for the secondary ray, and it is hard to believe that this could then give rise to numerous tertiary rays in the direction of the original beam fast enough to make their way into the Faraday cylinder. It seems possible that Glimme and Koenigsberger did not take sufficient precautions (by magnetic field or otherwise) to prevent the escape of δ-rays from the Faraday cylinder, and so got a spurious effect.

---

* 'Zeit. für Phys.,' p. 276 (1921).

Summary.

1. At a pressure of less than 1/100 mm. hydrogen positive rays of 10,000 volts mean energy suffer considerable small angle scattering in a distance of 15 cm.
2. This scattering is very much (10–20 times) greater than would be expected on theoretical grounds.
3. There must, therefore, be a field of force in the hydrogen molecule at distances of the order 10^{-8} from a nucleus, which is much stronger than would be expected from the inverse square law.
4. A subsidiary experiment throws great doubt on Glimme and Koenigsberger's "stossstrahlen."

In conclusion, I wish to express my very sincere thanks to Mr. W. S. Farren, M.A., Trinity College, for designing and making the slit mechanism shown in fig. 2 and for his help in preparing the illustrations. I must also record my deep gratitude to my father, Sir J. J. Thomson, O.M., for his valuable advice on many occasions during the progress of these researches.

Multenions and Differential Invariants.—II.

By ALEX. MCAULAY, Professor of Mathematics in the University of Tasmania.

(Communicated by W. B. Hardy, Sec.R.S. Received June 22, 1921.)

§ 11. *The fundamental covariant vector linity*, η. The meaning of a manifold as based on a quadratic differential form is here taken for granted, and this, whether, in the ordinary terminology, all or only some of its dimensions are real. Let

$$\left.\begin{array}{ll} \iota_a = i_a, & (a = 1,\ 2,\ \ldots c) \\ \iota_b = i_b\sqrt{(-1)}, & (b = c+1,\ c+2,\ \ldots n) \\ \rho = \sum\limits_{h=1}^{n} x_h\iota_h & \end{array}\right\}, \tag{1}$$

c having any given value from 1 to n. In the present paper we shall deal with multenions and multenion linities based on a system of vectors ρ for which all the co-ordinates x_h and the similar scalars are real. In the case of relativity $n = 4$, $c = 3$.

The fundamental quadratic form is taken to be

$$ds^2 = \mathrm{V}_0 d\rho\eta d\rho = \mathrm{V}_0 d\rho^{\backprime}\eta^{\backprime}d\rho^{\backprime}, \tag{2}$$

the sign which is given to ds^2, a scalar, being convenient in applications to relativity, but inconvenient in applications to pure geometry. η, a self-conjugate vector linity, is a given function of position, ρ; $\rho^{\backprime}$, a vector, an arbitrary function of ρ; and since the transformation from η to $\eta^{\backprime}$ is given by (2) the fundamental linity η is said to be covariant. [η and ds^2 are real, but sometimes ds and sometimes $ds\sqrt{(-1)}$ is real; when ds is real it is "a time-like interval"; when $ds\sqrt{(-1)}$ is real it is a length; when the question of reality is left undecided ds is an interval.] At any given position $\rho^{\backprime}$ may be so chosen that $\eta^{\backprime}d\rho^{\backprime} = d\rho^{\backprime}$, that is $\eta^{\backprime} = 1$. This imposes that η is without singularities; it has n mutually orthogonal principal directions and all its roots are real and positive, excluding zero values. Thus $\eta^{\frac{1}{2}}$ is assumed to have the same orthogonal system and its roots are taken to be the positive values of the square roots of η so that $\eta^{\frac{1}{2}}$ is real and without ambiguity.

We may, and from here on shall, take our previously introduced scalar k (§ 8) to be $|\eta^{\frac{1}{2}}|$ the determinant of $\eta^{\frac{1}{2}}$. [We have

$$\eta = \chi'\eta^{\backprime}\chi$$

so that $$|\eta| = |\eta^{\backprime}| \times |\chi|^2 = |\eta^{\backprime}| \times (k/k^{\backprime})^2,$$

because $k/k^{\backprime}$ was defined by

$$kdb = k^{\backprime}db^{\backprime} \quad \text{or} \quad k/k^{\backprime} = |\chi|.]$$

It will be seen that when, in the usual notation of relativity, it is said that with Galilean co-ordinates

$$(g_{11}, g_{22}, g_{33}, g_{44}) = (-1, -1, -1, +1),$$

we have not that $\eta(\iota_1, \iota_2, \iota_3, \iota_4) = (-\iota_1, -\iota_2, -\iota_3, +\iota_4)$ but instead the simpler statement that $\eta = \eta^{i} = 1$ or

$$\eta(\iota_1, \iota_2, \iota_3, \iota_4) = (\iota_1, \iota_2, \iota_3, \iota_4),$$

for it will be observed that this gives

$$ds^2 = -dx_1^2 - dx_2^2 - dx_3^2 + dx_4^2.$$

A little care is required in expressing a linity, its conjugate, reciprocal, etc., by scalars or co-ordinates. It seems sufficient to illustrate from vector linities. When we say that ϕ is given by the square array of 16 scalars

$$\left|\begin{matrix} a & b & c & d \\ e & f & g & h \\ l & m & n & p \\ q & r & s & t \end{matrix}\right|$$

we imply that the co-ordinates of $\phi\iota_1$ are found in the first *column* thus

$$\phi\iota_1 = a\iota_1 + e\iota_2 + l\iota_3 + q\iota_4$$

and so for the other columns. In the square array of n^2 scalars specifying ϕ, let a_{xy} be the scalar in the *xth row and yth column.* Thus $\phi\iota_y = \sum_{x=1}^{n} a_{xy}\iota_x$, or more conveniently

$$a_{xy} = \mathrm{V}_0\iota_x^{-1}\phi\iota_y. \tag{3}$$

Except when $c = n$ the square array of a self-conjugate linity is not symmetrical about the principal diagonal. Let a_{xy}, a_{xy}', be the square arrays of ϕ and its conjugate ϕ'. By the meaning of conjugate we have

$$\mathrm{V}_0\iota_x\phi\iota_y = \mathrm{V}_0\iota_y\phi'\iota_x.$$

Thus instead of $a_{yx}' = a_{xy}$ (as we have when $c = n$) we have by (3)

$$a_{yx}' = \iota_x^2\iota_y^2 a_{xy}, \tag{4}$$

and in particular, when ϕ is self-conjugate $a_{yx} = \iota_x^2\iota_y^2 a_{xy}$.

If we would retain the 16 scalars g_{ab} with their usual meanings as in treatises on relativity we may use $g_{(ab)}$ for the corresponding scalars specifying η. Thus from the fact that $ds^2 = \Sigma g_{ab}dx_a dx_b$ and from (3)

$$g_{ab} = \mathrm{V}_0\iota_a\eta\iota_b, \qquad g_{(ab)} = \mathrm{V}_0\iota_a^{-1}\eta\iota_b,$$

whence

$$g_{ab} = \iota_a^2 g_{(ab)}. \tag{5}$$

More details with regard to these notations and the theory of tensors will be found below in § 17.

Note on Notations in the Present Paper.—There will be many exceptions below to the conventions here to be described. Especially do exceptions occur in very frequently recurring symbols. Thus none of the following receive the "cross" notation although they should do according to the conventions,

$$\lambda,\ \mathrm{E}_a,\ \bar{\mathrm{E}}_a,\ \mathrm{C}_\omega,\ \mathrm{C},$$

λ is a covariant vector, and of the other four C_ω, C are contramixt linities while E_a, $\bar{\mathrm{E}}_a$ are portions of contramixt linities. The explanations of the technical terms will be made as they come to be used below.

(1) Contravariant multenions will have no distinguishing mark. Covariant multenions will usually be distinguished by a "cross" (which, because of its position, is scarcely likely to be confused with a sign of multiplication), thus

$$\alpha^\times,\ u^\times,\ p^\times,\ q^\times,$$

read "alpha cross," etc.

(2) Covariant and comixt linities will have no distinguishing mark. Contravariant and contramixt linities will usually be distinguished by a cross.

(3) Contravariant and covariant linities will be denoted by Greek letters; contramixt and comixt linities by Roman capitals.

(4) $\alpha, \beta, \gamma, \alpha^\times, \beta^\times, \gamma^\times$, will generally denote "dummy" vectors. (Eddington uses "dummy" in an analogous sense for a positive integer.) Dummies are always (below, though not of necessity) treated as constants in that they are not affected by the differentiations of ∇.

The duty of a dummy contravariant or covariant multenion may be generally described as "to occupy, and thereby indicate, a situation into which an actual, previously defined, contravariant or covariant multenion, may legitimately be inserted."

§ 12. *The Invariant Forms of any Multenion Expression.*—η has been defined as a fundamental *vector* linity. It is also used for the corresponding extensive linity, a multenion linity; and similarly for $\eta^{\frac{1}{2}}$. Since in (2), of § 11, $ds^2 = (\eta^{\frac{1}{2}}d\rho)^2$, $\eta^{\frac{1}{2}}d\rho$ in our manifold takes the place of the $d\rho$ of multenions in a Galilean manifold. [In a Euclidean manifold all the dimensions are real. In a Galilean manifold some are imaginary.] More generally, at a given position of the manifold, $\eta^{\frac{1}{2}}q$ where q is a contravariant multenion, or $\eta^{-\frac{1}{2}}q^\times$ where $q^\times$ is a covariant multenion, takes the place of q in a Galilean manifold. $\eta^{\frac{1}{2}}q$ and $\eta^{-\frac{1}{2}}q^\times$ will be spoken of as the neutral forms of q and $q^\times$. Indeed, the three q (contravariant), ηq (covariant) and $\eta^{\frac{1}{2}}q$ (neutral), are to be regarded as only different mathematical representations of the *same* intrinsic property, of the manifold; just as in an ordinary dynamical system, the whole system of velocities or instead the whole system of momenta equally will serve to specify the instantaneous motion.

Corresponding to any multenion expression such as $V_a . pqV_b(rs)$ which has a Galilean interpretation there are always the three forms (p, q, r, s now being contravariant) in our manifold (contra., co., neut.).

$$\eta^{-\frac{1}{2}}V_a . [\eta^{\frac{1}{2}}p\eta^{\frac{1}{2}}q V_b(\eta^{\frac{1}{2}}r\eta^{\frac{1}{2}}s)], \qquad \eta^{\frac{1}{2}}V_a . [\ \ ], \text{ and } V_a . [\ \ ].$$

The first and second of these three can *always* be expressed in such a shape that only η and not $\eta^{\frac{1}{2}}$ occurs explicitly; though the expression may be very complicated as in the example chosen. It is these final forms from which $\eta^{\frac{1}{2}}$ has been explicity banished, that are regarded as invariantive.

To show this, first note as an example, that the product pqr can be first modified by expressing qr as a sum of terms such as $V_{a+b-2c} . V_aqV_br$ and then $p . qr$ can be similarly modified. We have then merely to find the invariant forms (two, namely, covariant and contravariant) of $V_{a+b-2c}uv$. [The reader is reminded that in our notation u, v, w, have homogeneities a, b, c, respectively, or have the forms

$$u = V_ap,\ v = V_bq,\ w = V_ar.]$$

Let a be greater or equal to b and let u, v be contravariant. Then most frequently all we require to note are the two cases $c = 0$, $c = b$.

$$\eta^{-\frac{1}{2}}V_{a+b}\eta^{\frac{1}{2}}u\eta^{\frac{1}{2}}v = V_{a+b}uv, \qquad \eta^{\frac{1}{2}}V_{a+b}\eta^{\frac{1}{2}}u\eta^{\frac{1}{2}}v = V_{a+b}\eta u\eta v,$$

$$\eta^{-\frac{1}{2}}V_{a-b}\eta^{\frac{1}{2}}u\eta^{\frac{1}{2}}v = V_{a-b}u\eta v, \qquad \eta^{\frac{1}{2}}V_{a-b}\eta^{\frac{1}{2}}u\eta^{\frac{1}{2}}v = V_{a-b}\eta uv,$$

and here of course we may have $u = \eta^{-1}u^{\times}$ or $v = \eta^{-1}v^{\times}$ in each of the forms.

To deal with the general value of c in $V_{a+b-2c}uv$ we use ζ. Now we know that if $\tau_1, \tau_2, \ldots \tau_n$ are n independent contravariant vectors at a point, then their normal reciprocals $\bar{\tau}_1, \bar{\tau}_2, \ldots \bar{\tau}_n$ are n independent covariant vectors. Also we know that if $(\alpha, \beta^{\times})$ is any expression—bilinear in two vectors

$$(\zeta, \zeta) = -\sum_{a=1}^{n}(\tau_a, \bar{\tau}_a).$$

Hence if we insert ζ, ζ into two places in an invariant form, of which one is occupied by a contravariant vector and the other by a covariant vector, the resulting expression will be invariantive (meaning, "of invariant form"). Now it is easy to see by reduction to primitive units that

$$V_{a+b-2c}uv = [(-)^{\frac{1}{2}c(c+1)}/c!]\,V_{a+b-2c} . V_{a-c}u\zeta_c^{(c)}V_{b-c}\zeta_c^{(c)}v, \qquad (1)$$

where as usual $\zeta^{(c)}$ stands for $\zeta_1\zeta_2\ldots\zeta_c$ and $\zeta_c^{(c)}$ for $V_c\zeta^{(c)}$.

Hence the contravariant form of $V_{a+b-2c}uv$ is

$$[(-)^{\frac{1}{2}c(c+1)}/c!]\,V_{a+b-2c} . V_{a-c}u\zeta_c^{(c)}V_{b-c}\eta\zeta_c^{(c)}v, \qquad (2)$$

from which the covariant form may be written down, and either $u^{\times} = \eta u$ or $v^{\times} = \eta v$ or both may be used in place of u and v. In my MS. work I am in the habit of using $\zeta_K{}^{(c)}$, $\zeta_Q{}^{(c)}$ for $K\zeta_c{}^{(c)}$, $Q\zeta_c{}^{(c)}$. Thus (2) would be $(c\,!)^{-1}V_{a+b-2c}\,.\,V_{a-c}u\zeta_c{}^{(c)}V_{b-c}\eta\zeta_K{}^{(c)}v$. A special case of (2) often required is that the contravariant form of $V_2\omega\omega'$ where ω and ω' are contravariant and ${}_nV_2$ is

$$V_2(V_1\zeta\omega)(V_1\eta\zeta\omega'), \tag{3}$$

or more generally, for infinitesimal rotations generally, the contravariant form of $V_a\omega u$ is

$$V_a(V_1\zeta\omega)(V_{a-1}\eta\zeta u). \tag{4}$$

§ 13. *Normal and Incident Components in Invariant Form.*—These components occur very frequently in relativity. Thus the energy-momentum linity is density multiplied by an incident component, at any rate in its kinetic or primitive form. Normal and incident components serve to separate: (1) vector potential from electrostatic potential; (2) momentum from energy; (3) force from power; (4) current density from charge density; (5) magnetic induction from electrostatic force; (6) magnetic field strength from displacement. One should always be prepared to recognise these resolutions.

In quaternions if α is a given vector and σ an arbitrary one, σ has two components with reference to α, namely the normal component

$$N_0\sigma = \alpha^{-1}V\alpha\sigma = \alpha V\alpha\sigma\ .\ /\alpha^2$$

and the incident component

$$I_0\sigma = \alpha^{-1}S\alpha\sigma = \alpha S\alpha\sigma\ .\ /\alpha^2.$$

As will be familiar to users of the quaternion method, it is these components each multiplied by α^2, rather than the components themselves which we shall expect to meet with frequently in analytical expressions. The two

$$N_\alpha\sigma = N\sigma = \alpha V\alpha\sigma, \qquad I_\alpha\sigma = I\sigma = \alpha S\alpha\sigma,$$

may be referred to as the normal and incident *expressions* of σ with reference to α. With regard to these four linities N_0, I_0, N, I we have as follows

$$\left.\begin{array}{lll} N_0{}^2 = N_0, & I_0{}^2 = I_0, & N_0I_0 = 0 = I_0N_0 \\ N_0' = N_0, & I_0' = I_0, & N_0+I_0 = 1 = N_0'+I_0' \end{array}\right\}, \tag{1}$$

$$\left.\begin{array}{lll} N^2 = \alpha^2N, & I^2 = \alpha^2I, & NI = 0 = IN \\ N' = N, & I' = I, & N+I = \alpha^2 = N'+I' \end{array}\right\}. \tag{2}$$

Similarly in a Galilean manifold any multenion q has two components N_0q, I_0q normal and incident to a given vector α. Let

$$q = \sum_{c=1}^{n} w, \quad \text{where} \quad w = V_cq;$$

then when the linities Nw, N_0w, etc., have been defined, Nq, N_0q have also been defined. We have

(Galilean manifold.)
$$\left.\begin{array}{ll} Nw = \alpha V_{c+1}\alpha w, & Iw = \alpha V_{c-1}\alpha w \\ N_0 w = \alpha^{-1}V_{c+1}\alpha w, & I_0 w = \alpha^{-1}V_{c-1}\alpha w \end{array}\right\}. \tag{3}$$

[If $\acute{\epsilon}$ is a vectorium of homogeneity b, a vector σ has a normal component $\acute{\epsilon}^{-1}V_{b+1}\acute{\epsilon}\sigma$ and an incident component $\acute{\epsilon}^{-1}V_{b-1}\acute{\epsilon}\sigma$ with reference to $\acute{\epsilon}$. With reference to $\acute{\epsilon}$, w has many, not merely two, components given by

$$w = \acute{\epsilon}^{-1}V_{b+c}\acute{\epsilon}w + \acute{\epsilon}^{-1}V_{b+c-2}\acute{\epsilon}w + \ldots.$$

The terms on the right of this equation are all, separately, but not obviously, of homogeneity c.]

Next let α be a given contravariant vector at a given position ρ of our manifold and q an arbitrary contravariant multenion at the same position and let $w = V_c q$. Also let α_ν, q_ν, w_ν be the neutral forms and $\alpha^\times$, $q^\times$, $w^\times$ the covariant forms of α, q, w respectively, that is to say,

$$\left.\begin{array}{l} \eta^{\frac{1}{2}}\alpha = \alpha_\nu = \eta^{-\frac{1}{2}}\alpha^\times \\ \eta^{\frac{1}{2}}w = w_\nu = \eta^{-\frac{1}{2}}w^\times \\ \eta^{\frac{1}{2}}q = q_\nu = \eta^{-\frac{1}{2}}q^\times \end{array}\right\}. \tag{4}$$

The invariant form of α^2 is α_ν^2 or

$$V_0\alpha\eta\alpha = \alpha_\nu^2 = V_0\alpha^\times\eta^{-1}\alpha^\times. \tag{5}$$

With our present meanings (3) is not invariant in form. There are four modes of choosing the invariant forms of N and I, N_0 and I_0 and generally of any multenion linities, (1) covariant, (2) contravariant, (3) comixt, (4) contramixt. We will take number (3) of these for the basis of the notation in connection with N and I. A comixt linity means one which acting on a covariant multenion produces a covariant multenion. For this type of linity it is obvious that the invariant forms of (3) are given by

$$\left.\begin{array}{l} Nw^\times = V_c\alpha V_{c+1}\eta\alpha w^\times = V_c\eta^{-1}\alpha^\times V_{c+1}\alpha^\times w^\times \\ Iw^\times = V_c\eta\alpha V_{c-1}\alpha w^\times = V_c\alpha^\times V_{c-1}\eta^{-1}\alpha^\times w^\times \\ N_0 w^\times = Nw^\times / V_0\alpha\eta\alpha = Nw^\times / V_0\alpha^\times\eta^{-1}\alpha^\times \\ I_0 w^\times = Iw^\times / V_0\alpha\eta\alpha = Iw^\times / V_0\alpha^\times\eta^{-1}\alpha^\times \end{array}\right\}. \tag{6}$$

(2) comes with this notation to be modified in two ways only; α^2 is replaced by $V_0\alpha\eta\alpha$ or $V_0\alpha^\times\eta^{-1}\alpha^\times$ and in place of N′ and I′ we have a new kind of conjugate $\eta N'\eta^{-1}$, $\eta I'\eta^{-1}$ involving the fundamental linity η. Thus we now have

$$\left.\begin{array}{lll} N^2 = V_0\alpha\eta\alpha \,.\, N, & I^2 = V_0\alpha\eta\alpha \,.\, I, & NI = 0 = IN \\ \eta N'\eta^{-1} = N, & \eta I'\eta^{-1} = I, & N + I = V_0\alpha\eta\alpha = \eta(N' + I')\eta^{-1} \end{array}\right\}. \tag{7}$$

[Note that if we put $\eta\phi'\eta^{-1} = \phi_c$ we obtain for this new conjugate the familiar properties

$$(\phi+\psi)_c = \phi_c+\psi_c, \qquad (\phi\psi)_c = \psi_c\phi_c, \qquad (\phi_c)_c = \phi.]$$

The types of linity we have to recognise are given in the following short list.

Linity Types.

Name	Co-variant.	Contra-variant.	Co-mixt.	Contra-mixt.	Neutral.
Symbol	ξ	$\xi^\times$	X	$X^\times$	$X_\nu = \phi$
Meaning	$p^\times = \xi q$	$p = \xi^\times q^\times$	$p^\times = Xq^\times$	$p = X^\times q$	$p_\nu = X_\nu q_\nu$

Thus ξ converts a q (contravariant) into a $p^\times$ (covariant) and similarly for the other types.

The two mixed linities convert a multenion of a given type to one of the same type; the other two convert a given type to the *other* type. [In our present method the mixed linities could scarcely be more unhappily named.]

Corresponding to any one of the types we have, of course, by means of η, symbols of the other types with the same intrinsic significance. Thus, corresponding to the neutral form ϕ we have

$$\left.\begin{aligned} \xi &= \eta^{\frac{1}{2}}\phi\eta^{\frac{1}{2}}, & \xi^\times &= \eta^{-\frac{1}{2}}\phi\eta^{-\frac{1}{2}} \\ X &= \eta^{\frac{1}{2}}\phi\eta^{-\frac{1}{2}}, & X^\times &= \eta^{-\frac{1}{2}}\phi\eta^{\frac{1}{2}} \end{aligned}\right\}. \tag{8}$$

From the list and from (8) the following statements, in which ξ_1, ξ_2 mean two linities of the type ξ, etc., are easily established :—

(1) $\xi_1\xi_2$ has no invariant significance, nor has $\xi_1{}^\times\xi_2{}^\times$.

(2) X_1X_2 is of type X and $X_1{}^\times X_2{}^\times$ is of type $X^\times$.

(3) Although by (1), for integral values of c, ξ^c has in general no invariant significance; in particular when $c = -1$ it has significance. ξ^{-1} is of type $\xi^\times$, and $\xi^{\times-1}$ is of type ξ. X^{-1} is of type X, and $X^{\times-1}$ is of type $X^\times$.

(4) A rational function $f(X)$ is of type X; and $f(X^\times)$ is of type $X^\times$; the corresponding neutral form in each case being $f(\phi)$.

(5) ξ' is of type ξ, $\xi^{\times\prime}$ of type $\xi^\times$; so that a similar remark applies to the self-conjugate and skew parts of ξ and $\xi^\times$. ξ' and $\xi^{\times\prime}$ correspond to ϕ'.

(6) X' is of type $X^\times$, $X^{\times\prime}$ of type X, both corresponding to ϕ'. Again $\eta X'\eta^{-1}$ is of type X and $\eta^{-1}X^{\times\prime}\eta$ of type $X^\times$, both corresponding to ϕ'.

(7) The identity satisfied by ξ or by $\xi^\times$ has no significance. The

identity satisfied by ϕ is the same as the identity satisfied by each of the following

$$X = \xi\eta^{-1} = \eta\xi^{\times}, \qquad X^{\times} = \eta^{-1}\xi = \xi^{\times}\eta,$$

and all their conjugates and η-conjugates.

(8) In the neutral form a finite rotation is effected by a vector linity (and its extensive) ϕ for which $\phi'\phi = 1$. This is equivalent to

$$\xi'\eta^{-1}\xi\eta^{-1} = \eta X'\eta^{-1}X = 1.$$

From the above it will be seen that if in a Galilean manifold we would deal with ϕ', $\phi' \pm \phi$, etc., then probably in the general manifold it is best to deal with ξ and $\xi^{\times}$; if in the Galilean we would use $\phi^{a}, f(\phi)$ etc., then in the general manifold the use of X, $X^{\times}$ is indicated.

We need not write down the twenty-four alternative forms of the eight equations (6), but we may note how in terms of N, I the other types appear.

$$\left.\begin{array}{l} N,\ I \text{ are of type } X \\ N' = \eta^{-1}N\eta,\quad I' = \eta^{-1}I\eta \text{ are of type } X^{\times} \\ N\eta,\ I\eta \text{ are of type } \xi \\ \eta^{-1}N,\ \eta^{-1}I \text{ are of type } \xi^{\times} \end{array}\right\}. \tag{9}$$

After the fundamental linity η, of type ξ, has presented itself the other three may be regarded as arising from

$$\xi^{-1}, \qquad \xi_1\xi_2^{-1}, \qquad \xi_1^{-1}\xi_2.$$

§ 14. *Riemann's and Weyl's Manifolds.*—Suppose ρ, $\rho+\alpha$, $\rho+\beta$ are for our manifold the position vectors of three infinitesimally near points, P, A, B. We assume that for such small regions the metrics are intrinsically the same as in a Galilean system, and in particular that there is a definite shift of PA along PB, and of PB along PA, each called a translation, by which if A in the first shift comes to C, B also in the second shift will come to C. With arbitrary co-ordinates we cannot of course affirm that C will have the position vector $\rho+\alpha+\beta$, but instead that it will have

$$\rho+\alpha+\beta+d_{\alpha}^{\tau}\beta$$

where d_{α}^{τ} is to be regarded, not as a symbol of operation, but merely as the abbreviation of the words

$$d_{\alpha}^{\tau} = \text{“ increment due to translation along } \alpha\text{.”} \tag{1}$$

From these preliminaries it is clear that

$$d_{\alpha}^{\tau}\beta = d_{\beta}^{\tau}\alpha = -E_{\alpha}\beta, \tag{2}$$

where E_{α} is a vector linity which is a function of the position vector ρ, and is also a function of α, linear in the latter; also $E_{\alpha}\beta$ is symmetrical in α and β.

The word translation implies something about constancy or otherwise in interval or in "the square measure" ds^2 of equation (2) of § 11. In Weyl's view we are to assume that in general *a translation is such* that it is impossible for the square measure to remain unaltered when by successive steps a closed path has been traced. In such a closed path beginning and ending at a given point, there will appear a change in the square measure depending on the point and the path.

The square measure then is assumed to change according to the equation

$$d_\alpha{}^{\mathrm{T}} \mathrm{V}_0 \beta\eta\beta = -\mathrm{V}_0\beta\eta\beta \,.\, \mathrm{V}_0 \lambda\alpha, \tag{3}$$

where α and β are arbitrary and with the same meanings as before, and λ is a vector, a given function of ρ. Thus α, β must be contravariant and λ covariant. $\mathrm{E}_\alpha\beta$ is a vector neither covariant nor contravariant, which need not surprise us, because more than one position of the manifold has been used in its description.

It is clear that η is arbitrary as to a scalar factor and λ correspondingly arbitrary. More definitely, nothing is altered in the meanings given if (1) η is multiplied by a positive scalar h, any function of position, and (2) λ is altered to $\lambda + \nabla \log h$, as appears from equation (3).

Important qualification of the last sentence.—η here means the original *vector* linity denoted by η. Whenever we speak of altering gauge by multiplying η by h this restriction is to be understood. It is clear that ηw thereby changes, not to $h\eta w$, but to $h^c\eta w$.

"Into all expressions or formulæ, which exhibit analytically true metrical properties of the manifold, η and λ must enter in such a manner that invariance subsists (1) when there is arbitrary transformation of co-ordinates (ρ replaced by ρ'); and (2) when η, λ are replaced by $h\eta$, $\lambda + \nabla \log h$, whatever positive scalar function of ρ, h may be."*

Since (3) is true, and also what (3) becomes both when β is changed to γ and when changed to $\beta+\gamma$, we get the corresponding bilinear form

$$d_\alpha{}^{\mathrm{T}} \mathrm{V}_0 \beta\eta\gamma = -\mathrm{V}_0\beta\eta\gamma \,.\, \mathrm{V}_0 \lambda\alpha, \tag{4}$$

where α, β, γ are all arbitrary. In the translation denoted by $d_\alpha{}^{\mathrm{T}}$, η suffers the increment $-\mathrm{V}_0\alpha\nabla \,.\, \eta$ and β and γ the increments $-\mathrm{E}_\alpha\beta$, $-\mathrm{E}_\alpha\gamma$. Hence

$$-\mathrm{V}_0\alpha\nabla_\vartheta \,.\, \mathrm{V}_0\beta\eta_\vartheta\gamma - \mathrm{V}_0\mathrm{F}_\alpha\beta\eta\gamma - \mathrm{V}_0\beta\eta\mathrm{E}_\alpha\gamma = -\mathrm{V}_0\beta\eta\gamma \,.\, \mathrm{V}_0\lambda\alpha.$$

Put

$$\eta\mathrm{E}_\alpha = \Gamma_\alpha, \tag{5}$$

so that $\Gamma_\alpha\beta$, like $\mathrm{E}_\alpha\beta$, is symmetrical in α, β. Thus

$$-\mathrm{V}_0(\gamma\Gamma_\alpha\beta + \beta\Gamma_\alpha\gamma) = \mathrm{V}_0\alpha(\nabla_\vartheta - \lambda) \,.\, \mathrm{V}_0\beta\eta_\vartheta\gamma.$$

---

* Weyl, 'Raum, Zeit, Materie,' 3rd ed., p. 110. English translation, 1922, of 4th ed. (Brose), last sentence of p. 123.

Write down the two equations obtained by cyclically changing α, β, γ, change the signs of the second and third, and then add the three equations. We obtain

$$2V_0\alpha\Gamma_\beta\gamma = V_0\alpha\,[(\acute{\nabla}_9-\lambda)\,V_0\beta\eta_9\gamma - V_0\beta(\nabla_9-\lambda)\,.\,\eta_9\gamma - V_0\gamma(\nabla_9-\lambda)\,.\,\eta_9\beta],$$

or

$$\Gamma_\beta\gamma = \tfrac{1}{2}[-V_0\beta(\nabla_9-\lambda)\,.\,\eta_9\gamma - V_0\gamma(\nabla_9-\lambda)\,.\,\eta_9\beta + (\nabla_9-\lambda)\,V_0\beta\eta_9\gamma]. \quad (6)$$

For Riemann's manifold we have merely to put $\lambda = 0$. Since frequent reference must be made to this case of Riemann's manifold, we will put for the values of E_α and Γ_α when $\lambda = 0$, $\bar{E}_\alpha$ and $\bar{\Gamma}_\alpha$ respectively.

From (6)

$$(\Gamma_\beta + \Gamma_\beta')\gamma = -V_0\beta(\nabla_9-\lambda)\,.\,\eta_9\gamma, \quad (7)$$

$$(\Gamma_\beta - \Gamma_\beta')\gamma = V_1[V_2(\nabla_9-\lambda)\,\eta_9\beta\,.\,\gamma]. \quad (8)$$

Thus $\Gamma_\beta'\gamma$ is obtained from (6) by reversing the signs of the second and third terms on the right. It further follows that $\Gamma_\beta'\gamma$ regarded as a linity of β, instead of γ, is self-conjugate.

This, and its converse, can be shown to be true of any such function $Y_\alpha\beta$ which is symmetrical in α and β. Thus we may put for the ι_c component of $Y_\alpha\beta$, $\iota_c V_0\alpha\psi_c\beta$ where ψ_c is a self-conjugate vector linity, or

$$\left.\begin{aligned} Y_\alpha\beta &= \sum_{c=1}^{n} \iota_c V_0\beta\psi_c\alpha \\ \text{whence}\qquad Y_\alpha'\beta &= \sum_{c=1}^{n} \psi_c\alpha V_0\iota\beta \end{aligned}\right\}, \quad (9)$$

showing that $Y_\alpha'\beta$ is self-conjugate with respect to α. Each ψ_c involves $\frac{1}{2}n(n+1)$ scalars, so that $Y_\alpha\beta$ involves $\frac{1}{2}n^2(n+1)$ scalars; when $n = 4$ this number is 40.

When $Y_\alpha\beta$ is not symmetrical (9) still holds, ψ_c now being not self-conjugate, and the number of scalars being n^3.

We shall from this point onwards take for granted, as is not difficult to prove, that *at a given point it is always possible so to choose our co-ordinates that all the* $\frac{1}{2}n^2(n+1)$ *differential coefficients* (∇_9, η_9) *are zero; and independently to choose the gauge* (h) *so that* λ *is zero; therefore also to make* E_α *zero for all values of* α. [Brose calls h the calibration ratio. Selection of a definite h he calls calibration.]

§ 15. *Absolute Differentiation.*—Let τ be a contravariant vector given over an assigned path or over a region of n or fewer dimensions. If the region is of fewer than n dimensions, we may still assign arbitrary values over a region of n dimensions containing the given region, so that we can continue to apply ∇ to it.

In passing from the point ρ to the point $\rho+\alpha$, where α is infinitesimal, the *ordinary* increment $-V_0\alpha\nabla . \tau$ will be denoted by $d\alpha^0\tau$, and the excess of this above $d_\alpha{}^T\tau$ will be called the absolute increment and denoted by $d_\alpha{}^A\tau$. Thus we have

$$\left.\begin{aligned} d_\alpha{}^A\tau &= d_\alpha{}^0\tau - d_\alpha{}^T\tau = d_\alpha{}^0\tau + E_\alpha\tau \\ &= -V_0\alpha\nabla . \tau + E_\alpha\tau \end{aligned}\right\}. \tag{1}$$

We are about to show, as might be expected, that this absolute increment, $d_\alpha{}^A\tau$, is contravariant. If

$$\chi = -V_0(\ )\nabla . \rho', \tag{2}$$

and the change from ρ to ρ' causes the multenion function q of position to change to q', and the multenion linity ϕ to change to ϕ'; then the necessary and sufficient conditions that q should be (1) contravariant, (2) covariant, are

$$(1)\ q' = \chi q, \qquad (2)\ q' = \chi'^{-1}q.$$

The necessary and sufficient conditions that ϕ should be (1) covariant, (2) contravariant, (3) comixt, (4) contramixt, are that ϕ' should be equal to

$$(1)\ \chi'^{-1}\phi\chi^{-1}, \qquad (2)\ \chi\phi\chi', \qquad (3)\ \chi'^{-1}\phi\chi', \qquad (4)\ \chi\phi\chi^{-1}.$$

All these may be put in infinitesimal forms which are generally easier to apply. Let

$$\rho' = \rho + \delta\rho = \rho + \epsilon, \qquad \chi' = 1 + \delta\chi = 1 + \chi_0,$$
$$q' = q + \delta q, \qquad \phi' = \phi + \delta\phi.$$

Then the above six conditions (necessary and sufficient) for the several types of invariance become

$$(1)\ \delta q = \chi_0 q, \qquad (2)\ \delta q = -\chi_0' q,$$

and in the case of ϕ, $\delta\phi$ is equal to

$$(1)\ -\chi_0'\phi - \phi\chi_0, \quad (2)\ \chi_0\phi + \phi\chi_0', \quad (3)\ -\chi_0'\phi + \phi\chi_0', \quad (4)\ \chi_0\phi - \phi\chi_0;$$

and in order to make the tests we have

$$\chi_0 = -V_0(\ )\nabla . \epsilon, \qquad \chi_0' = -\nabla_2 V_0(\ )\epsilon_2. \tag{3}$$

[χ is applied to general multenions by the rules applicable to extensives; χ_0 by those applicable to subextensives.]

First apply the test to $E_\alpha\tau$; is $\delta E_\alpha\tau = \chi_0 E_\alpha\tau$? Instead of working this directly it will be found easier first to find $\delta V_0\alpha\Gamma_\beta\gamma$ from the equation preceding (6) of § 14, in which observe that, since $V_0\alpha\nabla = V_0\alpha'\nabla'$, δ has no effect on $V_0\alpha\nabla$, $V_0\beta\nabla$, $V_0\gamma\nabla$. Further for α, β, γ,

$$\delta(\alpha, \beta, \gamma) = \chi_0(\alpha, \beta, \gamma),$$

and for η

$$\delta\eta = -\chi_0'\eta - \eta\chi_0.$$

It is thus easy to prove that

$$\delta V_0\alpha\Gamma_\beta\gamma = -V_0\beta\nabla_2 . V_0\gamma\nabla_2 . V_0\eta\alpha\epsilon_2,$$

and thence that

$$\delta E_\beta \gamma = \chi_0 E_\beta \gamma - V_0 \beta \nabla \, . \, V_0 \gamma \nabla \, . \, \epsilon, \tag{4}$$

so that $E_a \tau$ is not contravariant. Next we have

$$\begin{aligned}\delta \, . \, d_a{}^0 \tau - \chi_0 d_a{}^0 \tau &= -\delta (V_0 \alpha \nabla \, . \, \tau) + \chi_0 (V_0 \alpha \nabla \, . \, \tau) \\ &= -V_0 \alpha \nabla \, . \, \chi_0 \tau + \chi_0 (V_0 \alpha \nabla \, . \, \tau)\end{aligned}$$

because δ does not affect $V_0 \alpha \nabla$, and $\delta \tau = \chi_0 \tau$. Thus

$$\delta \, . \, d_a{}^0 \tau = \chi_0 d_a{}^0 \tau + V_0 \alpha \nabla_\vartheta \, . \, V_0 \tau \nabla_\vartheta \, . \, \epsilon_\vartheta$$

whence finally

$$\delta d_a{}^{\mathrm{A}} \tau = \chi_0 d_a{}^{\mathrm{A}} \tau,$$

or $d_a{}^{\mathrm{A}} \tau$ is contravariant.

In one respect our taking of equation (1) to define $d_a{}^{\mathrm{A}} \tau$ has been misleading. If we follow on from it and equation (4) § 14 in a natural manner to define $d_a{}^{\mathrm{A}} \sigma^\times$ where $\sigma^\times$ is a covariant vector, the meaning arrived at will be on the whole inconvenient. For the moment we may take it as a sufficient reason for this statement that if equation (4) § 14 is to hold for any scalar invariant function of position, x, then $d_a{}^{\mathrm{T}} x$ being different from zero, if $d_a{}^{\mathrm{A}} x$ is to be equal to $d_a{}^0 x - d_a{}^{\mathrm{T}} x$, absolute and ordinary increments of invariants must differ. We reject the suggestion of equation (1), then, except for the increment of a *contravariant* vector. It is convenient to introduce a new increment, the invariantive increment, $d_a{}^{\mathrm{I}} q$. For τ we will take

$$d_a{}^{\mathrm{I}} \tau = d_a{}^{\mathrm{T}} \tau = -E_a \tau, \tag{5}$$

but to fix the meaning for other cases we will take that

$$d_a{}^{\mathrm{I}} x = 0 \tag{6}$$

when x is an invariant scalar function of position. This gives

$$0 = d_a{}^{\mathrm{I}} V_0 \sigma^\times \tau = V_0 (d_a{}^{\mathrm{I}} \sigma^\times) \tau - V_0 \sigma^\times E_a \tau.$$

Since τ is arbitrary we get

$$d_a{}^{\mathrm{I}} \sigma^\times = E_a{}' \sigma^\times. \tag{7}$$

Finally in all cases we put

$$d_a{}^{\mathrm{A}} = d_a{}^0 - d_a{}^{\mathrm{I}} = -V_0 \alpha \nabla \, . \, - d_a{}^{\mathrm{I}}. \tag{8}$$

The strongest reason for this meaning of $d_a{}^{\mathrm{A}}$ remains to be given. It is the only increment of the kind which is *truly* invariantive for *truly* contravariant and covariant vectors; vectors which besides their usual relations to change of co-ordinates are also invariant for change of gauge. If τ is thus truly contravariant $\eta \tau$ cannot be truly covariant, and therefore we ought not to base our conception of the standard increment of $\sigma^\times$ on supposing it possible to put $\sigma^\times = \eta \tau$.

Nevertheless the following up of the meaning of $d_a{}^{\mathrm{T}} \sigma^\times$ from equation (4) of § 14 has its uses, and we shall make one important application of it when later we consider curvature. It is not difficult to show that

$$d_a^{\mathrm{T}}\sigma^{\times} = \mathrm{E}_a'\sigma^{\times} - \sigma^{\times}\mathrm{V}_0\alpha\lambda, \qquad d_a^{\mathrm{T}}\eta = d_a^{0}\eta = -\mathrm{V}_0\alpha\nabla \,.\, \eta \tag{9}$$

(so that if we had continued to take $d_a^{\blacktriangle} = d_a^{0} - d_a^{\mathrm{T}}$ we should have found, as in the Riemann manifold, that $d_a^{\blacktriangle}\eta = 0$). With the meaning of (8) above $d_a^{\blacktriangle}\eta$ is not zero, as in a Riemann manifold, but instead

$$d_a^{\blacktriangle}\eta = -\mathrm{V}_0\alpha\lambda \,.\, \eta$$

when η is used in its primitive sense as a vector linity. When used as an extensive linity in ηw we get

$$d_a^{\blacktriangle}\eta \,.\, w = -c\mathrm{V}_0\alpha\lambda \,.\, \eta w. \tag{10}$$

We will now find the absolute increments of the two kinds of multenions and the four of multenion linities. Mixed linities interpreted as subextensive, have invariant meanings, but not the other two types. Thus if ϕ is contramixt

$$\phi\mathrm{V}_3\alpha\beta\gamma = \mathrm{V}_3\phi\alpha\beta\gamma + \mathrm{V}_3\alpha\phi\beta\gamma + \mathrm{V}_3\alpha\beta\phi\gamma$$

is of invariant form, because then $\phi\alpha$, β, γ are of one type, and is not of such form when ϕ is covariant, because then $\phi\alpha$, β, γ are not all of one type.

Let E_a and E_a' have the subextensive meanings as well as their primitive vector linity meanings. Thus

$$\mathrm{E}_a w = -\mathrm{V}_c\mathrm{E}_a\zeta\mathrm{V}_{c-1}\zeta w, \qquad \mathrm{E}_a' w = -\mathrm{V}_c\mathrm{E}_a'\zeta\mathrm{V}_{c-1}\zeta w. \tag{11}$$

We at once have

$$d_a^{\blacktriangle}p = d_a^{0}p - d_a^{\mathrm{T}}p = -\mathrm{V}_0\alpha\nabla \,.\, p + \mathrm{E}_a p, \tag{12}$$

$$d_a^{\blacktriangle}q^{\times} = d_a^{0}q^{\times} - d_a^{\mathrm{T}}q^{\times} = -\mathrm{V}_0\alpha\nabla \,.\, q^{\times} - \mathrm{E}_a'q^{\times}. \tag{13}$$

For the four linities $\xi_{\times}$, $\xi^{\times}$, X, $\mathrm{X}^{\times}$, treat of the first as an example of all four. The meaning of $d_a^{\blacktriangle}\xi$ is given by

$$d_a^{\blacktriangle}\xi \,.\, p = d_a^{\blacktriangle}(\xi p) - \xi d_a^{\blacktriangle}p.$$

Here ξp is a $q^{\times}$ so that

$$d_a^{\blacktriangle}(\xi p) = -\mathrm{V}_0\alpha\nabla \,.\, (\xi p) - \mathrm{E}_a'\xi p.$$

Hence $$d_a^{\blacktriangle}\xi \,.\, p = -\mathrm{V}_0\alpha\nabla_9 \,.\, \xi_9 p - \mathrm{E}_a'\xi p - \xi\mathrm{E}_a p.$$

Thus

$$\left.\begin{aligned} d_a^{\blacktriangle}\xi &= -\mathrm{V}_0\alpha\nabla \,.\, \xi - \mathrm{E}_a'\xi - \xi\mathrm{E}_a \\ d_a^{\blacktriangle}\xi^{\times} &= -\mathrm{V}_0\alpha\nabla \,.\, \xi^{\times} + \mathrm{E}_a\xi^{\times} + \xi^{\times}\mathrm{E}_a' \\ d_a^{\blacktriangle}\mathrm{X} &= -\mathrm{V}_0\alpha\nabla \,.\, \mathrm{X} - \mathrm{E}_a'\mathrm{X} + \mathrm{X}\mathrm{E}_a' \\ d_a^{\blacktriangle}\mathrm{X}^{\times} &= -\mathrm{V}_0\alpha\nabla \,.\, \mathrm{X}^{\times} + \mathrm{E}_a\mathrm{X}^{\times} - \mathrm{X}^{\times}\mathrm{E}_a \end{aligned}\right\}. \tag{14}$$

It is instructive to re-write in the present notation the four kinds of $\delta\phi$ above

$$\left.\begin{aligned}\delta\xi &= -\chi_0'\xi-\xi\chi_0\\ \delta\xi^\times &= \chi_0\xi^\times+\xi^\times\chi_0'\\ \delta X &= -\chi_0'X+X\chi_0'\\ \delta X^\times &= \chi_0X^\times-X^\times\chi_0\end{aligned}\right\}. \qquad (15)$$

Doubtless the complete parallelism between (14) and (15) has some simple explanation. Of the four results in each set, the second follows from the first by putting $\xi^\times = \xi^{-1}$ and the fourth from the third by putting $X^\times = X'$.

We obtain six very general formulæ showing how to convert such expressions as $(\nabla_\theta, q_\theta^\times)$ or $(\nabla_\theta, p_\theta)$ to invariant form [provided $(\alpha^\times, q^\times)$ or $(\alpha^\times, p)$ are already bilinear and invariantive] by one common process. Introduce a second notation for $d_\alpha^\wedge$, thus

$$d_\alpha^\wedge q = q_\alpha, \qquad d_\alpha^\wedge\phi = \phi_\alpha, \text{ etc.}$$

Put ζ in place of α in (12), (13), and (14). Thus

$$(\zeta, p_\zeta) = (\nabla_\theta, p_\theta)+(\zeta, E_\zeta p), \qquad (12a)$$

$$(\zeta, q_\zeta^\times) = (\nabla_\theta, q_\theta^\times)-(\zeta, E_\zeta' q^\times), \qquad (13a)$$

$$\left.\begin{aligned}(\zeta, \xi_\zeta) &= (\nabla_\theta, \xi_\theta)+(\zeta, -E_\zeta'\xi-\xi E_\zeta)\\ (\zeta, \xi_\zeta^\times) &= (\nabla_\theta, \xi_\theta^\times)+(\zeta, E_\zeta\xi^\times+\xi^\times E_\zeta')\\ (\zeta, X_\zeta) &= (\nabla_\theta, X_\theta)+(\zeta, -E_\zeta'X+XE_\zeta')\\ (\zeta, X_\zeta^\times) &= (\nabla_\theta, X_\theta^\times)+(\zeta, E_\zeta X^\times-X^\times E_\zeta)\end{aligned}\right\}. \qquad (14a)$$

(12), (13), (14) are particular cases of (12*a*), (13*a*), (14*a*).

As an example of the use of these equations let us suppose we have with Galilean co-ordinates established the equation $T_\theta\nabla_\theta = 0$ where T is the (complete) energy momentum linity; a linity which in the general manifold is best taken as comixt; a self-conjugate *vector* linity (self-conjugacy means for a *comixt* linity $T = \eta T'\eta^{-1}$, that is $\eta^{-1}T$ or $T\eta$ is self-conjugate in the ordinary sense). The (truly) invariant form of $T_\theta\nabla_\theta$ is given as $T_\zeta\zeta$ by the third of (14*a*). For generality we will at first not assume T to be self-conjugate. We have

$$T_\zeta\zeta = T_\theta\nabla_\theta-E_\zeta'T\zeta+TE_\zeta'\zeta.$$

To evaluate $E_\zeta'T\zeta$, write down $E_\alpha'\beta^\times = \Gamma_\alpha'\eta^{-1}\beta^\times$ from (6) of § 14, change α, $\beta^\times$ to ζ, $T\zeta$, and simplify. Thus

$$E_\zeta'T\zeta = \tfrac{1}{2}\eta_\theta(\eta^{-1}T-T'\eta^{-1})(\nabla_\theta-\lambda)-\tfrac{1}{2}(\nabla_\theta-\lambda)\,V_0\zeta\eta_\theta\eta^{-1}T\zeta. \qquad (16)$$

When η^{-1} T is self-conjugate we get

$$E_\zeta'T\zeta = -\tfrac{1}{2}(\nabla_\theta-\lambda)\,V_0\zeta\eta_\theta\eta^{-1}T\zeta. \qquad (17)$$

Further, since $|\phi|^{-1} . d|\phi| = -V_0\zeta\phi^{-1}d\phi\zeta$, putting $T = 1$ we get

$$E_\zeta'\zeta = k^{-1}\nabla k - \tfrac{1}{2}n\lambda. \tag{18}$$

In equation (16), the skew part of $\eta^{-1}T$ alone occurs in the first term, and the self-conjugate part alone in the second. $\eta^{-1}T$ is contravariant. Let us put $\eta^{-1}T = \xi^\times$ and use the second of (14*a*) for the further reduction

$$\xi_\zeta^\times\zeta = \xi_\vartheta^\times\nabla_\vartheta + E_\zeta\xi^\times\zeta + \xi^\times E_\zeta'\zeta.$$

We may use (18) to modify $E_\zeta'\zeta$. $E_\zeta\xi^\times\zeta$ happens to be particularly simple as it involves only the self-conjugate part of $\xi^\times$; let $\xi^\times = \xi_0^\times + \xi_1^\times$ where $\xi_0^\times$ is the self-conjugate part and $\xi_1^\times$ the skew part. We find that

$$E_\zeta\xi^\times\zeta = \eta^{-1}(\eta_\vartheta\xi_0^\times + \tfrac{1}{2}V_0\zeta\eta_\vartheta\xi_0^\times\zeta)(\nabla_\vartheta - \lambda) \tag{17a}$$

and $$\xi_\zeta^\times\zeta = k^{-1}(k\xi^\times)_\vartheta(\nabla_\vartheta - \tfrac{1}{2}n\lambda) + \eta^{-1}(\eta_\vartheta\xi_0^\times + \tfrac{1}{2}V_0\zeta\eta_\vartheta\xi_0^\times\zeta)(\nabla_\vartheta - \lambda). \tag{17b}$$

Linities of the form $\phi + \frac{1}{2}V_0\zeta\phi\zeta$ frequently present themselves in relativity, ϕ being mixed and generally self-conjugate in the mixed sense.

Note that $\xi_\zeta\zeta$ and $X_\zeta^\times\zeta$ have no invariant meaning because $\xi\beta^\times$ and $X^\times\beta^\times$ have none. We thus do not seek generalisations of $\xi_\vartheta\nabla_\vartheta$ and $X_\vartheta^\times\nabla_\vartheta$. ∇ is a symbolic *covariant* vector because $V_0d\rho\nabla = V_0d\rho'\nabla'$.

(17*a*) is found again below and numbered (22), in the course of illustrating a generally useful process.

Returning to the generalisation required of $T_\vartheta\nabla_\vartheta = 0$, when T is comixt and η^{-1} T is self-conjugate, we get

$$(kT)_\vartheta(\nabla_\vartheta - \tfrac{1}{2}n\lambda) + \tfrac{1}{2}(\nabla_\vartheta - \lambda)V_0\zeta\eta_\vartheta\eta^{-1}(kT)\zeta = 0. \tag{19}$$

It is clear that kT, a vector linity *density* as it is called, rather than T itself, should be regarded as the fundamental function. (19) is our standard form on the hypothesis that T is *truly* invariantive. If it were of power a in the sense about to be explained, the first ∇_ϑ in (19) should be replaced by $\nabla_\vartheta - a\lambda$.

If an invariantive symbol has such a meaning that on change of gauge ($h\eta$ for η, $\lambda + h^{-1}\nabla h$ for λ) the symbol is multiplied by h^a then it is said to be of power a. η is of power 1 when used in its primitive sense as a vector linity; in ηw, η is of power c; k is of power $\frac{1}{2}n$. It is because of these powers that we find when ∇ applies to η (power 1) as in E_a and Γ_a, ∇ appears only in the combination $\nabla - \lambda$; when ∇ applies to k (power $\frac{1}{2}n$) we always find it in the combination $\nabla - \frac{1}{2}n\lambda$. Thus in (18), although in a Riemann manifold we should find $E_\zeta'\zeta = k^{-1}\nabla k$, we must expect in the Weyl manifold to find instead

$$E_\zeta'\zeta = k^{-1}(\nabla - \tfrac{1}{2}n\lambda)k.$$

The reason for this is easily seen. Thus if q is of power a, $(\nabla_\vartheta, q_\vartheta)$ is of no definite power at all, but as is quite easy to verify $(\nabla_\vartheta - a\lambda, q_\vartheta)$ is of power a. To be *truly* invariantive a symbol must be of power zero.

An important case of this kind occurs in connection with the earliest invariant expressions we met, namely,

$$k^{-1}\mathrm{V}_{a-1}\nabla(ku), \qquad \mathrm{V}_{b+1}\nabla v^\times.$$

Our standard forms for these are based on the supposition that u and $v^\times$ are each of power zero so that $\mathrm{V}_{b+1}\nabla v^\times$ has not to be modified; but in the other case ku is of power $\frac{1}{2}n$ and our new invariant form should be $k^{-1}\mathrm{V}_{a-1}(\nabla - \frac{1}{2}n\lambda)(ku)$. Now (12$a$) and (13$a$) show us that our standard forms in these two cases should be

$$\mathrm{V}_{a-1}\nabla u + \mathrm{V}_{a-1}\zeta\mathrm{E}_\zeta u, \qquad \mathrm{V}_{b+1}\nabla v^\times - \mathrm{V}_{b+1}\zeta\mathrm{E}_\zeta' v^\times.$$

These must be the same as those we have just been considering so that we must have

$$\mathrm{V}_{a-1}\zeta\mathrm{E}_\zeta u = k^{-1}\mathrm{V}_{a-1}[(\nabla - \tfrac{1}{2}n\lambda)k\,.\,u], \tag{20}$$

$$\mathrm{V}_{b+1}\zeta\mathrm{E}_\zeta' v^\times = 0. \tag{21}$$

These identities can be established directly, but not very easily.

The specially simple form of $\Gamma_\alpha + \Gamma_\lambda'$ given in (7) of § 14 furnishes a second useful result from (16) for it gives

$$(\Gamma_\zeta' + \Gamma_\zeta)\eta^{-1}\mathrm{T}\zeta = \eta_\vartheta\eta^{-1}\mathrm{T}(\nabla_\vartheta - \lambda)$$

where $\Gamma_\zeta'\eta^{-1} = \mathrm{E}_\zeta'$. Thus from (16)

$$\Gamma_\zeta\eta^{-1}\mathrm{T}\zeta = -\tfrac{1}{2}\eta_\vartheta(\eta^{-1}\mathrm{T} + \mathrm{T}'\eta^{-1})(\nabla_\vartheta - \lambda) + \tfrac{1}{2}(\nabla_\vartheta - \lambda)\mathrm{V}_0\zeta\eta_\vartheta\eta^{-1}\mathrm{T}\zeta. \tag{22}$$

First we have the specially simple case $\mathrm{T} = 1$, namely

$$\Gamma_\zeta\eta^{-1}\zeta = \eta_\vartheta\eta^{-1}\nabla_\vartheta - k^{-1}\nabla k + \tfrac{1}{2}(n-2)\lambda. \tag{23}$$

Another very simple particular case is when $\eta^{-1}\mathrm{T}$ is skew. The right of (22) is then zero, and $\eta^{-1}\mathrm{T}$ is of the form $\mathrm{V}_1\omega(\quad)$ where ω is a contravariant ${}_n\mathrm{V}_2$. Thus

$$\Gamma_\zeta\mathrm{V}_1\omega\zeta = 0.$$

Since $\Gamma_\alpha = \eta\mathrm{E}_\alpha$ we also have $\mathrm{E}_\zeta\mathrm{V}_1\omega\zeta = 0$. This last may be regarded as a special case of (21) and suggests two other forms of (20), (21). Operate on (20) by $\mathrm{V}_0 w^\times(\quad)$ taking $c = a-1$; and on (21) by $\mathrm{V}_0 w(\quad)$ taking $c = b+1$. We thus get

$$\mathrm{E}_\zeta'\mathrm{V}_{c+1}w^\times\zeta = k^{-1}\mathrm{V}_{c+1}w^\times(\nabla - \tfrac{1}{2}n\lambda)k, \tag{24}$$

$$\mathrm{E}_\zeta\mathrm{V}_{c-1}w\zeta = 0, \tag{25}$$

(18) is a particular case of (24).

§ 16. *Curvature.*—If (α, β) is any function of two vectors, linear in each then

$$(\alpha, \beta) - (\beta, \alpha) = (\zeta, \dot{\mathrm{V}}_1[\mathrm{V}_2\alpha\beta\,.\,\zeta]) \tag{1}$$

as is seen by putting $(\alpha, \beta) = -(\zeta, \beta V_0 \alpha\zeta)$ and $(\beta, \alpha) = -(\zeta, \alpha V_0 \beta\zeta)$. Thus when

$$(\alpha, \beta) = -(\beta, \alpha) \quad \text{then} \quad (\alpha, \beta) = \tfrac{1}{2}(\zeta, V_1[V_2\alpha\beta \,.\, \zeta]). \tag{2}$$

The application of these two equations should be constantly borne in mind in what follows, even when attention is not called to it.

Similarly for the various invariant forms of infinitesimal rotations of a vector,

$$V_1\omega\alpha^{\times}, \qquad V_1\omega^{\times}\alpha, \qquad V_1\omega\eta\alpha, \qquad V_1\omega^{\times}\eta^{-1}\alpha^{\times}.$$

Let us calculate the increase in γ, a contravariant vector, due to translation, starting from the point ρ, along the four vector sides of an infinitesimal parallelogram in the order α, β, $-\alpha$, $-\beta$, α and β being contravariant. Afterwards let us calculate the similar increments of $\eta\gamma$. We shall put $V_2\alpha\beta = \omega$, and the increments due to translation round the parallelogram will be denoted by $d_\omega{}^{\mathrm{T}}\gamma$, $d_\omega{}^{\mathrm{T}}\eta\gamma$. Since the difference of two contravariant vectors at a point ρ is itself contravariant it follows that $d_\omega{}^{\mathrm{T}}\gamma$ is contravariant, and similarly $d_\omega{}^{\mathrm{T}}\eta\gamma$ is covariant; and it is obvious that they must both be linear in both ω and γ. We shall, therefore, put

$$d_\omega{}^{\mathrm{T}}\eta\gamma = \eta d_\omega{}^{\mathrm{T}}\gamma = \eta C_\omega\gamma = \Omega_\omega\gamma, \tag{3}$$

C_ω being a contramixt vector linity and Ω_ω the corresponding covariant linity, which may be called the curvature linities. Along the side $+\alpha$ the increment in γ is $-E_\alpha\gamma$. Along the side $-\alpha$, γ has changed to $\gamma - E_\beta\gamma$ and $d_{-\alpha}{}^{\mathrm{T}} = -d_\alpha{}^{\mathrm{T}}$ has received the increment $-V_0\beta\nabla.(\quad)$. Thus, these two sides contribute increments

$$\begin{aligned} &-E_\alpha\gamma + (E_\alpha - V_0\beta\nabla \,.\, E_\alpha)(\gamma - E_\beta\gamma),\\ &= -V_0\beta\nabla \,.\, E_\alpha\gamma - E_\alpha E_\beta\gamma. \end{aligned}$$

The contributions of the other two sides are obtained by interchange of α and β and reversal of the sign. Hence, noting that E_α is linear in α,

$$C_\omega\gamma = d_\omega{}^{\mathrm{T}}\gamma = -E_{\varrho\,.\,V_1\omega\nabla_\varrho}\gamma + (-E_\alpha E_\beta + E_\beta E_\alpha)\gamma. \tag{4}$$

[Since $d_\alpha{}^{\mathrm{A}} = d_\alpha{}^{0} - d_\alpha{}^{\mathrm{T}}$, and since $d_\alpha{}^{0}$ taken round the parallelogram produces zero we get that $C_\omega\tau = \tau_{\alpha\beta} - \tau_{\beta\alpha}$ which shows that absolute differentiations in different directions are not commutative.]

In treating of $\eta\gamma$ in place of γ, instead of $-E_\alpha$ we have to write $E_\alpha{}' - V_0\lambda\alpha$ for $d_\alpha{}^{\mathrm{T}}$. Noting that if x, y are two scalars $(-V_0\lambda\alpha, -V_0\lambda\beta)$

$$-(E_\alpha{}' + x)(E_\beta{}' + y) + (E_\beta{}' + y)(E_\alpha{}' + x) = -E_\alpha{}'E_\beta{}' + E_\beta{}'E_\alpha{}'$$

we see at once how (4) has to be modified. Thus

$$\Omega_\omega\gamma = d_\omega{}^{\mathrm{T}}\eta\gamma = E_{\vartheta\,.\,V_1\omega\nabla_\vartheta}{}'\eta\gamma + (-E_\alpha{}'E_\beta{}' + E_\beta{}'E_\alpha{}')\eta\gamma - \eta\gamma \,.\, V_0\omega\omega^{\times}, \tag{5}$$

where

$$\omega^{\times} = V_2\nabla\lambda \tag{6}$$

[ω and $\omega^\times$ are of very different natures. ω is a contravariant *dummy*, $\omega^\times$ is a given truly covariant function of position; in relativity $\omega^\times$ is magnetic induction cum electrostatic intensity, λ being vector potential cum electrostatic potential.]

From (4) on putting $\omega = V_2\alpha\beta$, and noting that $E_\alpha\gamma$ is symmetrical in α and γ, the following "cyclic rule" is seen to be true for each of the two parts of $C_\omega\gamma$ separately

$$C_{V_2\alpha\beta}\gamma + C_{V_2\beta\gamma}\alpha + C_{V_2\gamma\alpha}\beta = 0. \tag{7}$$

Remembering that $\Omega_\omega = \eta C_\omega$, inspection of (4) and (5) shows that

$$\left.\begin{aligned} \Omega_\omega + \Omega_\omega' &= -V_0\omega\omega^\times \,.\, \eta \\ \text{or}\qquad C_\omega + \eta^{-1}C_\omega'\eta &= -V_0\omega\omega^\times \end{aligned}\right\}. \tag{8}$$

The equation in $\Omega_\omega + \Omega_\omega'$ is even simpler than that in $\Gamma_\alpha + \Gamma_\alpha'$; and for a Riemann manifold $\Omega_\omega + \Omega_\omega' = 0$.

The first of (8) shows that $\Omega_\omega = -\frac{1}{2}V_0\omega\omega^\times.\eta + V_1\omega'(\quad)$ where ω' is a covariant linear function of ω. Hence

$$\Omega_\omega\gamma = -\tfrac{1}{2}V_0\omega\omega^\times \,.\, \eta\gamma + V_1\varpi\omega\gamma, \tag{9}$$

where ϖ is a covariant ${}_nV_2$ linity. The first term on the right of (9) gives the ratio of stretch $-\frac{1}{2}V_0\omega\omega^\times$; the second term the rotation $\varpi\omega$; suffered by γ by $d_\omega^{\,T}$.

To find the properties of the linity ϖ first transform (5) to the following,

$$\Omega_\omega\gamma = \Gamma_{\vartheta\,.\,V_1\omega(\nabla_\vartheta-\lambda)}'\gamma + (\Gamma_\alpha'\eta^{-1}\Gamma_\beta - \Gamma_\beta'\eta^{-1}\Gamma_\alpha)\gamma - \eta\gamma \,.\, V_0\omega\omega^\times. \tag{10}$$

I will indicate in outline the mode of doing this. The first term on the right of (5) is

$$(\Gamma_\alpha'\eta^{-1})_\vartheta\,\eta\gamma V_0\beta\nabla_\vartheta - (\Gamma_\beta'\eta^{-1})_\vartheta\,\eta\gamma V_0\alpha\nabla_\vartheta$$

and the second term is

$$(-\Gamma_\alpha'\eta^{-1}\Gamma_\beta' + \Gamma_\beta'\eta^{-1}\Gamma_\alpha')\gamma = (\Gamma_\alpha'\eta^{-1}\Gamma_\beta - \Gamma_\beta'\eta^{-1}\Gamma_\alpha)\gamma + \text{two terms}$$

of which the first is $-\Gamma_\alpha'\eta^{-1}(\Gamma_\beta + \Gamma_\beta')\gamma$ and the second is $+\Gamma_\beta'\eta^{-1}(\Gamma_\alpha + \Gamma_\alpha')\gamma$. It will now be found that the parts depending on ∇ in $(\Gamma_\beta + \Gamma_\beta')$ and $(\Gamma_\alpha + \Gamma_\alpha')$ cancel with the following parts of the first term.

$$\Gamma_\alpha'(\eta^{-1})_\vartheta\eta\gamma \,.\, V_0\beta\nabla_\vartheta - \Gamma_\beta'(\eta^{-1})_\vartheta\eta\gamma \,.\, V_0\alpha\nabla_\vartheta.$$

Collecting the parts not cancelled we get (10).

In the first term and the third on the right of (10) there are terms of the form $(\nabla_\vartheta, \lambda_\vartheta)$. These give

$$\begin{aligned} &-\tfrac{1}{2}\eta\gamma \,.\, V_0\omega\omega^\times - \tfrac{1}{2}V_1 \,.\, V_2(\eta V_1\omega\nabla_\vartheta \,.\, \lambda_\vartheta)\gamma \\ &= -\tfrac{1}{2}\eta\gamma \,.\, V_0\omega\omega^\times - \tfrac{1}{4}V_1 \,.\, V_2(\eta V_1\omega\nabla_\vartheta \,.\, \lambda_\vartheta + \eta V_1\omega\lambda_\vartheta \,.\, \nabla_\vartheta)\gamma \\ &\qquad - \tfrac{1}{4}V_1 \,.\, V_2(\eta V_1\omega\nabla_\vartheta \,.\, \lambda_\vartheta - \eta V_1\omega\lambda_\vartheta \,.\, \nabla_\vartheta)\gamma. \end{aligned}$$

The last of these three parts furnishes the skew part ϖ_1, of ϖ, given by

$$\varpi_1\omega = \tfrac{1}{4}V_2(V_1\zeta\omega^\times \,.\, \eta V_1\zeta\omega). \tag{11}$$

The neutral form of this is the suggestive expression $\frac{1}{4}V_2\omega^\times\omega$. The remainder of ϖ, denoted by ϖ_0, the self-conjugate part, is

$$\left.\begin{aligned}\varpi_0\omega = &-\tfrac{1}{2}V_2[(\nabla_\vartheta-\lambda)\,\eta_\vartheta V_1\omega\,(\nabla_\vartheta-\lambda)]+\tfrac{1}{2}V_2\zeta_0\Gamma_\zeta'\eta^{-1}\Gamma_{\zeta_0}V_1\omega\zeta \\ &-\tfrac{1}{4}V_2(\eta V_1\omega\nabla_\vartheta\,.\,\lambda_\vartheta+\eta V_1\omega\lambda_\vartheta\,.\,\nabla_\vartheta)\end{aligned}\right\}. \tag{12}$$

So far we have divided $\Omega_\omega\gamma$ into three parts, such that the corresponding parts of $C_\omega\gamma$ are *truly* invariant, (1) the part self-conjugate in γ, namely, $-\frac{1}{2}\eta\gamma V_0\omega\omega^\times$; (2) the part $V_1\varpi_1\omega\gamma$ in which ϖ_1 is skew and (apart from the fundamental linity η) only involves $\omega^\times$ as with the first part; (3) the part $V_1\varpi_0\omega\gamma$ where ϖ_0 is self-conjugate. Part (3) obeys the cyclic rule; (1) and (2) together, but not separately, obey the same. The cyclic rule for ϖ_0 may be put, as will be shown later, in the form

$$V_4\,.\,V_2\zeta_1\zeta_2\varpi_0V_2\zeta_1\zeta_2=0,$$

so that ϖ_0 instead of having the full number

$$\tfrac{1}{2}\,.\,\tfrac{1}{2}n(n-1)\,.\,[\tfrac{1}{2}n(n-1)+1]$$

of scalars appropriate to a self-conjugate ${}_nV_2$ linity, has a smaller number because of the $n(n-1)(n-2)(n-3)/4\,!$ scalar equations of the cyclic condition just written. The number of scalars of ϖ_0 is thus $\frac{1}{12}n^2(n^2-1)$. For the parts (1) and (2) the $\frac{1}{2}n(n-1)$ scalars of $\omega^\times$ are required. The total number of scalars involved in C_ω is thus

$$\tfrac{1}{12}n(n-1)(n^2+n+6).$$

When $n=4$ the number is 26.

Putting $V_1\zeta\omega=\zeta_\omega$, $V_1\zeta\omega^\times=\zeta_{\omega\times}$ the two parts (1) and (2) together may be written as

$$\left.\begin{aligned}&-\tfrac{1}{4}\eta\gamma V_0\zeta_\omega\zeta_{\omega\times}+\tfrac{1}{4}V_1V_2(\zeta_{\omega\times}\eta\zeta_\omega)\gamma \\ &=\tfrac{1}{4}(-\eta\gamma\,.\,V_0\zeta_\omega\zeta_{\omega\times}-\eta\zeta_\omega\,.\,V_0\zeta_{\omega\times}\gamma+\zeta_{\omega\times}V_0\zeta_\omega\gamma\end{aligned}\right\}. \tag{13}$$

It is now desirable to separate ϖ_0 into three parts ϖ_a, ϖ_b, ϖ_c, which are covariant as to change of co-ordinates though not as to change of gauge, namely, the parts of degrees zero, one, two in λ.

ϖ_a, of course, gives precisely the Riemann terms, obtained by putting $\lambda=0$ in any expression for ϖ. Thus from (12)

$$\varpi_a\omega=-\tfrac{1}{2}V_2\,.\,\nabla_\vartheta\eta_\vartheta V_1\omega\nabla_\vartheta+\tfrac{1}{2}V_2\zeta_0\Gamma_\zeta'\eta^{-1}\Gamma_{\zeta_0}V_1\omega\zeta. \tag{14}$$

Next write down the terms involving differentiations of λ and make them invariantive by (13*a*) of § 15, using, however, $\bar{E}_\zeta'$ in place of E_ζ' as we desire invariance only for change of co-ordinates, and we desire *not* to introduce quadratic terms in λ. Thus we obtain for the probable value of ϖ_b

$$\left.\begin{aligned}\varpi_b\omega = &-\tfrac{1}{4}V_2(\eta V_1\omega\nabla_\vartheta\,.\,\lambda_\vartheta+\eta V_1\omega\lambda_\vartheta\,.\,\nabla_\vartheta) \\ &+\tfrac{1}{4}V_2(\eta V_1\omega\zeta\,.\,\bar{E}_\zeta'\lambda+\eta V_1\omega\bar{E}_\zeta'\lambda\,.\,\zeta)\end{aligned}\right\}. \tag{15}$$

Before verifying that this contains all the terms linear in λ, not differentiated, find the quadratic terms, $\varpi_c\omega$. Put Λ_a for the terms of Γ_a which contain λ, that is, $\Lambda_a = \Gamma_a - \overline{\Gamma}_a$. Thus

$$\varpi_c\omega = -\tfrac{1}{2} V_2\lambda\eta V_1\omega\lambda + \tfrac{1}{2} V_2\zeta_0\Lambda_\zeta{}'\eta^{-1}\Lambda_{\zeta_0}V_1\omega\zeta. \tag{16}$$

To simplify this put $\omega = V_2\alpha\beta$ and work in the neutral form, finally returning to invariant form. The result comes out that $4\varpi_c\omega$ is the normal expression of ω with reference to λ (in neutral form $\lambda V_3\lambda\omega$). Hence

$$\varpi_c\omega = \tfrac{1}{4}\eta V_2\lambda V_3\eta^{-1}\lambda\omega = \tfrac{1}{4} V_2\eta^{-1}\lambda V_3\lambda\eta\omega, \tag{17}$$

ϖ_a, ϖ_b, ϖ_c are all separately self-conjugate, and all separately obey the cyclic rule.

We will now show that (15) must contain all the terms of (12) which are linear in λ. Inspection of (12) shows that these terms form a function $(\nabla_\vartheta, \eta_\vartheta, \lambda, \omega)$ which is linear in all four constituents. If there is any term of this nature which has not been included in our three parts ϖ_a, ϖ_b, ϖ_c, such term must be covariant as to change of co-ordinates. Now $(\nabla_\vartheta, \eta_\vartheta, \lambda, \omega)$ can only be made covariant in this sense by the addition [first of (14*a*) of §15] to it of

$$(\zeta, -\overline{E}_\zeta{}'\eta - \eta\overline{E}_\zeta, \lambda, \omega) = -(\zeta, \overline{\Gamma}_\zeta{}' + \overline{\Gamma}_\zeta, \lambda, \omega) = -(\nabla_\vartheta, \eta_\vartheta, \lambda, \omega).$$

This addition reduces the supposed absent term to zero. Hence there is no such term.

We now proceed to the two successive "contractions" of Ω_ω or C_ω. The first contraction yields a vector linity denoted by Ω or C in its covariant or contramixt form. The second contraction yields a scalar, an invariant of course, F.

$\Omega_{V_2\zeta\beta}\eta^{-1}\zeta$ is a covariant vector because the subscript ζ occupies the place of a contravariant vector and the second ζ occupies the place of a covariant vector. Put then

$$\Omega\beta = \Omega_{V_2\zeta\beta}\eta^{-1}\zeta, \qquad C\beta = \eta^{-1}\Omega\beta = C_{V_2\zeta\beta}\eta^{-1}\zeta, \tag{18}$$

$$F = -V_0\zeta\eta^{-1}\Omega\zeta = -V_0\zeta C\zeta. \tag{19}$$

Note, that of the four symbols Ω_ω, C_ω, Ω, C, *the only two which are truly invariantive are* C_ω and Ω; Ω_ω is of power 1 and C of power -1; F also is of power -1.

Let us now obtain the terms contributed to $\Omega\beta$ and to F by our five parts of Ω_ω, that is, by the first term in (9) and the four parts of ϖ, namely, ϖ_1, ϖ_a, ϖ_b, ϖ_c. For the latter four other than ϖ_a work details in the neutral form, and at the last step put the result into invariant form by the rules of § 12. Where first differential coefficients of λ occur, attend in the detailed working only to the terms in which the differential coefficients occur, and in the last step use equation (13a), of § 15, modified by reading $\bar{E}_\zeta'$ instead of E_ζ' to obtain the invariant form.

Let us denote the parts contributed by the first term of equation (9), to Ω and F by Ω_2 and F_2, and those contributed by ϖ_1, ϖ_a, ϖ_b, ϖ_c, by Ω_1, Ω_a, Ω_b, Ω_c and by F_1, F_a, F_b, F_c, respectively. Thus we find

$$\left.\begin{aligned}\Omega_2\beta = -\tfrac{1}{2}V_1\omega^\times\beta, \qquad \Omega_1\beta = \tfrac{1}{4}(n-2)V_1\omega^\times\beta \\ (\Omega_2+\Omega_1)\beta = \tfrac{1}{4}(n-4)V_1\omega^\times\beta \qquad\qquad\end{aligned}\right\}. \tag{20}$$

Those are the only parts of $\Omega\beta$ that are not self-conjugate, and they are purely skew. It is a remarkable fact that Ω is self-conjugate for the case $n = 4$, and for no other value of n. Of course, this accidental circumstance, if we are to call it such, is fortunate for the theory of relativity. From (20), it at once follows that

$$F_2 = F_1 = 0. \tag{21}$$

It is easy to show that when a term ϖ_x of ϖ is self-conjugate the corresponding part of Ω_x of Ω is self-conjugate. ϖ_a, ϖ_b, ϖ_c, are self-conjugate. We find

$$\left.\begin{aligned}4\Omega_b\beta &= -2\eta\beta\,.\,k^{-1}V_0\nabla(k\eta^{-1}\lambda) - (n-2)(\nabla_\vartheta V_0\lambda_\vartheta\beta + \lambda_\vartheta V_0\nabla_\vartheta\beta + 2\bar{E}_\beta'\lambda) \\ F_b &= -(n-1)k^{-1}V_0\nabla(k\eta^{-1}\lambda)\end{aligned}\right\}. \tag{22}$$

In Ω_c the normal expression of β with respect to λ occurs thus:

$$\left.\begin{aligned}\Omega_c\beta &= \tfrac{1}{4}(n-2)V_1\eta^{-1}\lambda V_2\lambda\eta\beta \\ F_c &= \tfrac{1}{4}(n-1)(n-2)V_0\lambda\eta^{-1}\lambda\end{aligned}\right\}. \tag{23}$$

In a Riemann manifold ϖ_a completely specifies Ω_ω, and therefore Ω, $\Omega_{\omega a}$, and Ω_a are given by any expressions for Ω_ω and Ω by putting $\lambda = 0$. The most useful form of Ω_a is derived from (5). To get $\Omega\beta$ from (5) we have to put $\alpha = \zeta$, $\eta\gamma = \zeta$. Thus we obtain

$$\Omega\beta = E_{\beta\vartheta}'\nabla_\vartheta - E_\zeta'E_\beta'\zeta + E_\beta'E_\zeta'\zeta + V_0\beta\nabla\,.\,E_\zeta'\zeta + V_1\beta\omega^\times. \tag{24}$$

Here note that the first three terms are just those, when we put $E_\beta' = T$, that were under discussion in equations (16), (17), (18) of § 15, and (24) may be modified accordingly. Our immediate purpose, however, is to put λ, and therefore $\omega^\times$ zero, so as to obtain Ω_a. Noting that $\bar{E}_\zeta'\zeta = k^{-1}\nabla k$ by (18) of § 15, we have

$$\Omega_a\beta = \bar{E}_{\beta\vartheta}'\nabla_\vartheta - \bar{E}_\zeta'\bar{E}_\beta'\zeta + (\bar{E}_\beta' + V_0\beta\nabla\,.\,)\nabla\log k. \tag{25}$$

This is the form (26·3) given in Eddington's Report on Relativity, that is to say, Eddington's G_{bc} is our $V_0\iota_b\Omega_a\iota_c$ [see equation (7), § 17 below]. (25) gives

$$F_a = -V_0\eta^{-1}\zeta(\bar{E}_{\zeta\vartheta}'\nabla_\vartheta - \bar{E}_{\zeta_0}'\bar{E}_\zeta'\zeta_0 + [\bar{E}_\zeta' + V_0\zeta\nabla\,.\,])\nabla\log k. \tag{26}$$

Returning to equation (24) let us find a useful expression for $C_a\beta$. It is convenient on occasion to use the following notations

$$\eta^{-1}\Gamma_\beta\eta^{-1} = \Gamma_\beta^\times, \qquad \eta^{-1}\bar{\Gamma}_\beta\eta^{-1} = \bar{\Gamma}_\beta^\times,$$

but this is a little cumbrous for our present purpose. Let us put

$$\overline{\Gamma}_\beta{}^{\times\prime} = \eta^{-1}\overline{\mathrm{E}}_\beta{}' = \xi_\beta{}^\times = \mu_\beta{}^\times + \nu_\beta{}^\times, \tag{24a}$$

where $\mu_\beta{}^\times$ and $\nu_\beta{}^\times$ [see (24*c*) below] are put for the self-conjugate and skew parts of $\xi_\beta{}^\times$, and avoid the double suffix notation $\xi_{\beta 0}{}^\times$, $\xi_{\beta 1}{}^\times$. Putting $\lambda = 0$ in (24) we have

$$\mathrm{C}_a\beta = \eta^{-1}(\eta\xi_\beta{}^\times)_0\nabla_0 - \eta^{-1}\overline{\Gamma}_\zeta{}'\xi_\beta{}^\times\zeta + \xi_\beta{}^\times\nabla\log k + \eta^{-1}(\mathrm{V}_0\beta\nabla\,.\,\nabla\log k)$$

$$[\text{Put } -\overline{\Gamma}_\zeta{}' = \overline{\Gamma}_\zeta{}' - (\overline{\Gamma}_\zeta + \overline{\Gamma}_\zeta{}') = \overline{\Gamma}_\zeta + \mathrm{V}_0\zeta\nabla\,.\,\eta.]$$

$$= \xi_{\beta 0}{}^\times\nabla_0 + \overline{\mathrm{E}}_\zeta\xi_\beta{}^\times\zeta + k^{-1}\xi_\beta{}^\times\nabla k + \eta^{-1}(\mathrm{V}_0\beta\nabla\,.\,\nabla\log k),$$

whence, using (17*a*) of § 15 for $\overline{\mathrm{E}}_\zeta\xi_\beta{}^\times\zeta$, which involves $\mu_\beta{}^\times$ but not $\nu_\beta{}^\times$,

$$\mathrm{C}_a\beta = k^{-1}(k\xi_\beta{}^\times)_0\nabla_0 + \eta^{-1}(\eta_0\mu_\beta{}^\times + \tfrac{1}{2}\mathrm{V}_0\zeta\eta_0\mu_\beta{}^\times\zeta)\nabla_0 + \eta^{-1}(\mathrm{V}_0\beta\nabla\,.\,\nabla\log k). \tag{24b}$$

$\mu_\beta{}^\times$ and $\nu_\beta{}^\times$ being $\tfrac{1}{2}\eta^{-1}(\pm\Gamma_\beta + \Gamma_\beta{}')\eta^{-1}$ are written down from (7) and (8) of § 14. Thus

$$\left.\begin{aligned}\mu_\beta{}^\times\gamma^\times &= \tfrac{1}{2}\mathrm{V}_0\beta\nabla_\kappa\,.\,(\eta^{-1})_\kappa\gamma^\times\\ \nu_\beta{}^\times\gamma^\times &= -\tfrac{1}{2}\mathrm{V}_1[\eta^{-1}\mathrm{V}_2\nabla_\kappa\eta_\kappa\beta\,.\,\gamma^\times]\end{aligned}\right\}. \tag{24c}$$

From (24*b*) we may write down $k\mathrm{F}_a$ in the form $-k\mathrm{V}_0\zeta\mathrm{C}_a\zeta$ and we shall reach an expression useful for reducing, later, $\delta(k\mathrm{F}_a)$ resulting from changing η to $\eta + \delta\eta$.

In what immediately follows let us work with the neutral forms, only occasionally stopping to point out the slightly more complicated invariant forms. In the case of a vector linity, ϕ, we may say that there is a definite part, namely, $-n^{-1}\mathrm{V}_0\zeta\phi\zeta$, which gives rise to the contraction $-\mathrm{V}_0\zeta\phi\zeta$; the second part, $\phi + n^{-1}\mathrm{V}_0\zeta\phi\zeta$, has zero contraction and involves only $n^2 - 1$ scalars as against the full number, n^2, for a vector linity. Similarly, it is possible to separate out from any vector expression (ω, γ) linear in each constituent a definite part $(n-1)^{-1}(\mathrm{V}_2\zeta\mathrm{V}_1\omega\gamma, \zeta)$ absorbing (generally) n^2 scalars from the full number, $\tfrac{1}{2}n^3(n-1)$ required to specify (ω, γ). The remainder of (ω, γ), namely

$$(\omega, \gamma) - (n-1)^{-1}(\mathrm{V}_2\zeta\mathrm{V}_1\omega\gamma, \zeta)$$

has zero contraction. The part thus furnishing the contraction C of C_ω is

$$(n-1)^{-1}\mathrm{C}_{\mathrm{V}_2\zeta\mathrm{V}_1\omega\gamma}\zeta = (n-1)^{-1}\mathrm{C}\mathrm{V}_1\omega\gamma$$

and it obviously satisfies the cyclic rule (and would do so whether $\mathrm{C}_\omega\gamma$ itself did so or not).

The following form of the cyclic rule for $\mathrm{C}_\omega\gamma$ may be noticed:—

$$\mathrm{C}_{\mathrm{V}_2\zeta\mathrm{V}_3\omega\gamma}\zeta = 0. \tag{27}$$

This is true because by (1) of § 7

$$\mathrm{V}_2\zeta\mathrm{V}_3\omega\gamma = \mathrm{V}_2\beta\gamma\mathrm{V}_0\alpha\zeta + \mathrm{V}_2\gamma\alpha\mathrm{V}_0\beta\zeta + \mathrm{V}_2\alpha\beta\mathrm{V}_0\gamma\zeta$$

(27) is closely connected with a generalisation of (1) and (2) above. The cyclic rule is an identity which is combinatorial in α, β, γ; that is to say, the sign of the expression equated to zero alters by interchange of any two of the three α, β, γ. If $(\alpha_1 \alpha_2, \ldots, \alpha_a)$ is any expression linear in each vector and combinatorial in the vectors, it can always be expressed as a linear function of $\alpha_a^{(a)}$. It is actually equal to

$$(a\,!)^{-1}(\zeta_1, \zeta_2, \ldots, \zeta_{a-1}, \mathrm{V}_1 \alpha_a^{(a)} \zeta_{a-1}^{(a-1)}), \tag{28}$$

(28) may be proved by mathematical induction by proving the general connecting link in the following successive equalities, a being, for example, taken as equal to 4,

$$\begin{aligned}(\alpha, \beta, \gamma, \delta) &= (2\,!)^{-1}(\zeta, \mathrm{V}_1 \,.\, \mathrm{V}_2\alpha\beta \,.\, \zeta, \gamma, \delta)\\ &= (3\,!)^{-1}(\zeta_1, \zeta_2, \mathrm{V}_1 \,.\, \mathrm{V}_3\alpha\beta\gamma\mathrm{V}_2\zeta_1\zeta_2, \delta)\\ &= (4\,!)^{-1}(\zeta_1, \zeta_2, \zeta_3, \mathrm{V}_1 \,.\, \mathrm{V}_4\alpha\beta\gamma\delta\mathrm{V}_3\zeta_1\zeta_2\zeta_3).\end{aligned}$$

This is readily done when it is noticed that

$$\mathrm{V}_1\zeta_3''' \mathrm{V}_4\alpha\beta\gamma\delta = \Sigma \pm \alpha\mathrm{V}_0\zeta_3'''\beta\gamma\delta$$

and the like ($\zeta_3''' = \mathrm{V}_3\zeta_1\zeta_2\zeta_3$).

Since $$\mathrm{V}_2\zeta\mathrm{V}_1\omega\gamma - \mathrm{V}_2(\mathrm{V}_3\zeta\omega)\gamma = -\omega\mathrm{V}_0\zeta\gamma,$$

we have $$\mathrm{C}_\omega\gamma = \mathrm{C}\mathrm{V}_1\omega\gamma - \mathrm{C}_{\mathrm{V}_2(\mathrm{V}_3\zeta\omega)\gamma}\zeta \tag{29}$$

whether or not $\mathrm{C}_\omega\gamma$ satisfies the cyclic rule. If the rule is satisfied we have further

$$\mathrm{C}_\omega\gamma = \mathrm{C}_{\mathrm{V}_2(\mathrm{V}_1\zeta\omega)\gamma}\zeta. \tag{30}$$

This may be obtained from (29) or directly from the cyclic rule in its original form thus:

$$\mathrm{C}_\omega\gamma = \mathrm{C}_{\mathrm{V}_2\alpha\beta}\gamma = \mathrm{C}_{\mathrm{V}_2\alpha\gamma}\beta - \mathrm{C}_{\mathrm{V}_2\beta\gamma}\alpha = \mathrm{C}_{\mathrm{V}_2(\mathrm{V}_1\zeta\omega)\gamma}\zeta$$

by (1) above.

It is useful to modify the cyclic rule for a self-conjugate ϖ_j, namely, ϖ_0, ϖ_a, ϖ_b, or ϖ_c. Thus, using ϖ_a as an example, we have $\Sigma\mathrm{V}_1\alpha\varpi_a\mathrm{V}_2\beta\gamma = 0$, or operating by $\mathrm{V}_0\delta(\;)$, where δ is a fourth vector,

$$\mathrm{V}_0 \,.\, \mathrm{V}_2\alpha\beta\varpi_a\mathrm{V}_2\gamma\delta + \mathrm{V}_0 \,.\, \mathrm{V}_2\alpha\gamma\varpi_a\mathrm{V}_2\delta\beta + \mathrm{V}_0 \,.\, \mathrm{V}_2\alpha\delta\varpi_a\mathrm{V}_2\beta\gamma = 0.$$

The expression on the left is combinatorial in $\alpha, \beta, \gamma, \delta$. Taking $\alpha, \beta, \gamma, \delta$ to be any four of the primitive vectors $\iota_1, \iota_2, \ldots \iota_n$, say, $\iota_1, \iota_2, \iota_3, \iota_4$, we have successively the following three statements:—

$$\Sigma[\iota_1^{-1}\iota_2^{-1}\iota_3^{-1}\iota_4^{-1}\Sigma\mathrm{V}_0(\iota_1\iota_2)\varpi_a(\iota_3\iota_4)] = 0,$$

$$\mathrm{V}_4\zeta_1\zeta_2\zeta_3\zeta_4 \,.\, \mathrm{V}_0(\mathrm{V}_2\zeta_1\zeta_2)\varpi_a(\mathrm{V}_2\zeta_3\zeta_4) = 0.$$

$$\mathrm{V}_4(\mathrm{V}_2\zeta_1\zeta_2)\varpi_a(\mathrm{V}_2\zeta_1\zeta_2) = 0, \tag{31}$$

as mentioned above between equations (12) and (13).

From equation (27) onwards we have been at no pains to put our results in invariant form. Nevertheless (27), (30), and (31) are already in invariant form, and (29) yields

$$C_\omega\gamma = CV_1\omega\eta\gamma - C_{V_2(V_3\zeta\omega)\eta\gamma}\eta^{-1}\zeta. \tag{29a}$$

§ 17. *Comparison with Theory of Tensors.*—Let $\phi(q_1 q_2, \dots q_a)$ be any contravariant or covariant function linear in each of the multenions $q_1, q_2, \dots q_a$, each one of which is also contravariant or covariant, not necessarily all of one type. Thus, for instance, $\Omega_{V_2\alpha\beta}\eta^{-1}\gamma^\times$ is such a covariant function of $\alpha, \beta, \gamma^\times$, two contravariant vectors, and one covariant vector. In the pages above we have defined various increments of q due to changing ρ to $\rho+\alpha$, where α is an arbitrary infinitesimal contravariant vector

$$(d_\alpha^{\mathrm{T}},\ d_\alpha^{\mathrm{I}},\ d_\alpha^{\mathrm{O}},\ d_\alpha^{\Lambda})\, q$$

and we might add, when q is of power c, *the absolute increment proper*, also of power c, $d_\alpha^{\Lambda\Gamma}q = d_\alpha^{\mathrm{B}}q = (d_\alpha^{\Lambda} + cV_0\alpha\lambda)\, q$. The increments of ϕ of these types are all formed on the same principle. Thus take d_α^{Λ} as the type and use the abbreviated notation $d_\alpha^{\Lambda}q = q_\alpha$. Then the increment ϕ_α of ϕ is given by

$$\phi_\alpha(q_1, q_2, \dots q_a)$$

$$= [\phi(q_1, q_2, \dots q_a)]_\alpha - \sum_{c=1}^{a} \phi(q_1, q_2, \dots q_{c\alpha} \dots q_a)$$

where $q_{c\alpha}$ occupies the position of q_c in $\phi(q_1, q_2, \dots, q_a)$. Once having obtained our increment, we may cease to regard α as infinitesimal, so that d_α^{Λ} becomes, not so much an *increment* as a *rate of increase* multiplied by the magnitude of α, though above it has always been called an increment.

All this is on the model of the Theory of Tensors. In that theory, however, the only linear form $\phi(q_1, q_2, \dots, q_a)$ recognised is the scalar form $\phi(\alpha_1, \alpha_2, \dots, \alpha_n)$ in several vectors, some of which are covariant and the rest contravariant. The present writer marvels at the great results obtained by expounders of the Theory of Tensors, since the basis is such a limited concept.

The precise translation from our presentation to that of the Theory of Tensors and *vice versâ* is to a certain extent arbitrary, on account of vagueness, especially in the matter of sign, that writers on the Theory of Tensors seem to permit themselves. The following will be found to form one consistent system. Let A^a, B^b be two contravariant vectors, α and β, and let X_{ab} be a covariant tensor of the second rank, with the same intrinsic meaning as ξ, a covariant vector linity. A_aB_b seems to be taken to mean the same as B_bA_a in the Theory of Tensors, but I cannot see how this can be consistently done, and in the system here given A_aB_b and B_bA_a, when interpreted as vector linities, must have different meanings.

With the basis as given in § 11 above, we have as follows. Let

$$\alpha^{\times} = \eta\alpha, \qquad \beta^{\times} = \eta\beta, \tag{1}$$

$$\xi^{\times} = \eta^{-1}\xi\eta^{-1}, \qquad \mathrm{X} = \xi\eta^{-1}, \qquad \mathrm{X}^{\times} = \eta^{-1}\xi. \tag{2}$$

Then

$$\nabla = -\sum_{a=1}^{n} \iota_a{}^{-1}\frac{\partial}{\partial r_a}, \qquad \frac{\partial}{\partial r_a} = -\mathrm{V}_0\iota_a\nabla, \tag{3}$$

$$\mathrm{A}^a = \mathrm{V}_0\iota_a{}^{-1}\alpha, \qquad \mathbf{A}_a = \mathrm{V}_0\iota_a\alpha^{\times}, \tag{4}$$

$$g_{ab} = \mathrm{V}_0\iota_a\eta\iota_b, \qquad g^{ab} = \mathrm{V}_0\iota_a{}^{-1}\eta^{-1}\iota_b{}^{-1}, \qquad g_a{}^b = \mathrm{V}_0\iota_a\iota_b{}^{-1}, \tag{5}$$

$$[bc, a] = \mathrm{V}_0\iota_a\overline{\Gamma}_{\iota_b}\iota_c, \qquad \{bc, a\} = \mathrm{V}_0\iota_a{}^{-1}\overline{\mathrm{E}}_{\iota_b}\iota_c \tag{6}$$

$$\mathbf{X}_{ab} = \mathrm{V}_0\iota_a\xi\iota_b, \qquad \mathrm{X}_a{}^b = \mathrm{V}_0\iota_a\mathrm{X}\iota_b{}^{-1}, \qquad \mathbf{X}^{ab} = \mathrm{V}_0\iota_a{}^{-1}\xi^{\times}\iota_b{}^{-1}. \tag{7}$$

The equation $\mathrm{X}_{ab} = \mathrm{A}_a\mathrm{B}_b$ consistently with the above and the related equations may be interpreted to mean as follows

$$\mathrm{X}_{ab} = \mathrm{A}_a\mathrm{B}_b, \qquad \xi = \eta\alpha\mathrm{V}_0(\ )\eta\beta = \alpha^{\times}\mathrm{V}_0(\ )\beta^{\times}, \tag{8}$$

$$\mathrm{X}_a{}^b = \mathrm{A}_a\mathrm{B}^b, \qquad \mathrm{X} = \eta\alpha\mathrm{V}_0(\ )\beta = \alpha^{\times}\mathrm{V}_0(\ )\beta. \tag{9}$$

$$\mathrm{X}^{ab} = \mathrm{A}^a\mathrm{B}^b, \qquad \xi^{\times} = \alpha\mathrm{V}_0(\ )\beta. \tag{10}$$

Now (8) and (10) definitely tell us what $\mathrm{B}_b\mathrm{A}_a$ and $\mathrm{B}^b\mathrm{A}^a$ must mean; instead of representing ξ and $\xi^{\times}$ respectively, they represent ξ' and $\xi^{\times\prime}$ and $\mathrm{B}^b\mathbf{A}_a$ is similarly naturally interpreted as $\eta\beta\mathrm{V}_0()\alpha$, that is as $\eta\mathrm{X}'\eta^{-1}$.

Lastly we have

$$(\zeta, \zeta) = -\sum_{a=1}^{n} (\iota_a, \iota_a{}^{-1}). \tag{11}$$

It is with these interpretations that equation (25), § 16, was stated to be equivalent to Eddington's (26·3).

I may add here that Eddington's equation (23) for $\mathrm{B}_{abc}{}^h$ gives

$$\mathrm{B}_{abc}{}^h = -\mathrm{V}_0\iota_h{}^{-1}\overline{\mathrm{C}}_{\mathrm{V}_2\iota_b\iota_c}\iota_a \tag{12}$$

where C_ω means $\mathrm{C}_{\omega a}$, that is the part of C_ω obtained by putting $\lambda = 0$.

It will be seen that I interpret the mixed tensor to mean a comixt, not contramixt, vector linity.

18. *Prof. Eddington's recent work.** [Re-written September, 1921.]

Suppose ABC is a triangle whose sides AB, BC, CA are geodesics. Let the point O of an infinitesimal lattice-work pass round the triangle, starting and ending at A, and in each stage of the journey let the lattice-work suffer "parallel displacement" only. When the circuit has been completed, then:—according to Euclid and Galileo the lattice-work will be found occupying its initial position exactly; according to Riemann it will be found turned through an angle in a plane containing O; according to Weyl it will be turned and changed in size, but not in shape; according to Eddington it will be found turned, changed in size and also in shape. If Prof. Eddington meant no more than this, his manifold, as will be shown below, would be mathematically equivalent to Weyl's, but he does mean more. In his primary manifold *there is no criterion for shape at all*, none to distinguish a square from any other parallelogram, or to distinguish a cube from any other parallelepiped. In Riemann's manifold the fundamental idea is that of the invariant quadratic form (squared interval), and "parallel displacement" is a mathematically *derived* idea. In Eddington's manifold the *fundamental* idea is that of parallel displacement, and an invariant quadratic form is a mathematically derived idea. When $n = 4$, Riemann's, Weyl's and Eddington's foundations require ten, thirteen (nine *ratios* of coefficients of quadratic form and four coefficients of a linear form) and forty scalars at a point to be given, respectively.

---

* 'Roy. Soc. Proc.,' A, vol. 99, p. 104 (1921).

Eddington then starts with $E_\alpha\beta (= E_\beta\alpha, - E_\alpha\beta$ meaning increment of β due to the parallel displacement α; α, β being contravariant vectors) but without any fundamental η specifying the intrinsic invariant $V_0 d\rho\eta d\rho$. *All our previous work and results, which are independent of η, or can be made so, continue to be valid in Eddington's manifold.* On looking back I only see one case where our old reasoning ceases to hold in the new circumstances. This is equation (4), § 15, which was proved by aid of η. But the equation is still true because the condition that τ, a contravariant vector, suffers the parallel displacement β is

$$-V_0\beta\nabla \,.\, \tau = -E_\beta\tau,$$

and this is an intrinsic condition independent of choice of coordinates. Infinitesimally change the coordinates as in § 15 [$\delta\tau = \chi_0\tau$, $\delta\beta = \chi_0\beta$, $\delta(V_0\beta\nabla) = 0$]. Thus

$$-V_0\beta\nabla \,.\, (\chi_0\tau) = -\delta(E_\beta\tau);$$

and here the left transforms to

$$-\chi_0(V_0\beta\nabla \,.\, \tau) - V_0\beta\nabla_\vartheta \,.\, \chi_{0\vartheta}\tau = -\chi_0 E_\beta\tau - V_0\beta\nabla_\vartheta \,.\, \chi_{0\vartheta}\tau$$

so that

$$\delta(E_\beta\tau) = \chi_0 E_\beta\tau + V_0\beta\nabla_\vartheta \,.\, \chi_{0\vartheta}\tau.$$

This is equation (4) § 15 with τ in place of γ. We may then make the very important use of that equation which we made before, namely, if τ is *any* contravariant vector,

$$d_\alpha{}^A\tau = -V_0\alpha\nabla \,.\, \tau + E_\alpha\tau$$

is contravariant and may therefore be appropriately called the ***absolute*** increment of τ. [Prof. Eddington does not show in his paper that absolute differentiation thus holds good in the "in-" sense in his manifold.]

Our chief disability from absence of a fundamental η seems to be that there is now found no place for the ideas associated with normal and incident components (and expressions) and with rotations. $V_{a+b}uv$ and $V_{a-b}uv$ have their invariantive equivalents, but apparently $V_{a+b-2c}uv$ in general has not. We can still speak of α and $\beta^\times$ being normal when $V_0\alpha\beta^\times = 0$ and this leaves us with (what we met with in § 10 before η was introduced) the ideas of normal and incident components with reference not to a single τ but with reference to τ_c, some one of n given independent contravariant vectors $\tau_1, \tau_2, \dots \tau_n$. In a word the normal reciprocals of $\tau_1, \tau_2, \dots \tau_n$ continue to have invariantive meanings.

For present purposes the contraction we selected in § 16 to pass from C_ω to Ω was unhappily chosen as it depended on η. We will here, therefore, change the contraction to another. The previous definition of $\Omega\beta$ was that it was obtained by setting α, γ equal to $\zeta, \eta^{-1}\zeta$ in $\eta C_{V_2\alpha\beta}\gamma$; our new meaning of $\Omega'\alpha$ is obtained by setting $\beta, \gamma^\times$ equal to ζ, ζ in $C_{V_2\alpha\beta}'\gamma^\times$. Thus

$$\left.\begin{aligned}\text{Previous } \Omega\beta &= \eta C_{V_2\zeta\beta}\eta^{-1}\zeta \\ \text{Present } \Omega'\alpha &= C_{V_2\alpha\zeta}'\zeta\end{aligned}\right\}. \tag{1}$$

In Weyl's manifold the change in meaning is slight. The old and new self-conjugate parts are the same, but whereas the previous skew part of $\Omega\beta$ was $\frac{1}{4}(n-4)V_1\omega^\times\beta$, the present is $-\frac{1}{4}nV_1\omega^\times\beta$. This simplicity of connections between old and new meanings may be proved from the identity [equation (8), § 16].

$$\eta C_{V_2\alpha\beta}\gamma + C_{V_2\alpha\beta}'\eta\gamma = -V_0\omega^\times V_2\alpha\beta\,.\,\eta\gamma,$$

by setting α, γ equal to $\zeta, \eta^{-1}\zeta$. The identity is the one which expresses that our lattice-work, after its journey round the circuit ABC, is changed in size but not in shape.

Our equation for C_ω (4) of § 16 still holds while η is unassigned, and from it

$$C_\omega'\gamma^\times = -E_{\vartheta\,.\,V_1\omega\nabla_\vartheta}'\gamma^\times + (-E_\beta'E_\alpha' + E_\alpha'E_\beta)\gamma^\times. \tag{2}$$

Putting here $V_1\omega\nabla = \alpha V_0\beta\nabla - \beta V_0\alpha\nabla$ and using (1) we get

$$\Omega'\alpha = E_{\vartheta.\alpha}'\nabla_\vartheta + V_0\alpha\nabla\,.\,E_\zeta'\zeta - E_\zeta E_\alpha'\zeta + E_\alpha'E_\zeta'\zeta. \tag{3}$$

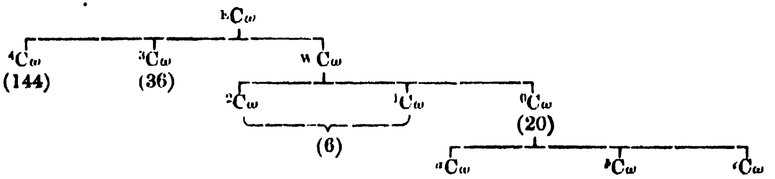
The first, third and fourth terms on the right are all self-conjugate because $E_\alpha\beta = E_\beta\alpha$. Hence the skew part of $\Omega\alpha$ is $-\frac{1}{2}V_1\alpha V_2\nabla E_\zeta'\zeta$.

By its original meaning $C_\omega\gamma$ is the difference of two truly contravariant, or as Eddington says in-contravariant (would not *preter*-contravariant, be better ?) vectors at the same point. Hence C_ω, C_ω', Ω and $V_2\nabla E_\zeta'\zeta$ all have the "in-" property and $\frac{1}{2}(\Omega + \Omega')$ is fitted for serving as the fundamental η of a Riemann manifold. Prof. Eddington, therefore, identifies

$\frac{1}{2}cV_0 d\rho(\Omega + \Omega')d\rho$ with Einstein's quadratic form, the squared interval; c being an absolute constant. I do not propose to follow him in this respect, but to show instead, that as soon as it is recognized that "lengths" can be compared (even in Weyl's limited sense that equality of lengths can be postulated *at* a point only), then Eddington's theory is mathematically equivalent to Weyl's; and instead of merely one simple quadratic form there are several available to identify with Einstein's.

It is now necessary to distinguish between Weyl's and Eddington's meanings of E_a, C_ω, Ω and related symbols. For Eddington's meanings we shall use $\mathbf{E}_a$, ${}^{E}C_\omega$, ${}^{E}\Omega$. For Weyl's we shall use E_a, ${}^{W}C_\omega$, ${}^{W}\Omega$.

The following three modifications of notations occurring in §§ 14 to 16 above are desirable: (1) The letter G will be used in place of the letter F, because F is wanted below for another purpose; (2) the "contraction" to pass from C_ω to Ω (both of power zero), will be altered in meaning as already described; (3) the symbols 2, 1, 0, a, b, c, used as suffixes to C_ω, Ω_ω, C, Ω, ϖ, G (former F), will be shifted from the right-hand bottom corner of the letter affected to the left-hand top corner. The following scheme will probably be helpful:—



[In each of these four lines there is one symbol equivalent to a tensor of Eddington's, or to a tensor of Weyl's. ${}^{E}C_\omega$ is equivalent to Eddington's ${}^{*}B_{abc}{}^{h}$, ${}^{W}C_\omega$ to Weyl's $F_{abc}{}^{h}$, ${}^{0}C_\omega$ to Weyl's ${}^{*}F_{abc}{}^{h}$, and ${}^{a}C_\omega$ to both Eddington's $B_{abc}{}^{h}$ and Weyl's $R_{abc}{}^{h}$.] In the scheme the E of ${}^{E}C_\omega$ implies "in Eddington's sense"; the W of ${}^{W}C_\omega$ implies "in Weyl's sense"; further, the scheme implies the following three defining equations:—

$$
\begin{aligned}
{}^{E}C_\omega &= {}^{4}C_\omega + {}^{3}C_\omega + {}^{W}C_\omega, \\
{}^{W}C_\omega &= {}^{2}C_\omega + {}^{1}C_\omega + {}^{0}C_\omega, \\
{}^{0}C_\omega &= {}^{a}C_\omega + {}^{b}C_\omega + {}^{c}C_\omega.
\end{aligned}
$$

We may write x for any one of the ten symbols E, W, 4, 3, 2, 1, 0, a, b, c. For seven out of the ten meanings of ${}^{x}C_\omega$, namely, for all except ${}^{a}C_\omega$, ${}^{b}C_\omega$, ${}^{c}C_\omega$, it is true that ${}^{x}C_\omega$ is of zero power as meant by Weyl. The three ${}^{a}C_\omega$, ${}^{b}C_\omega$, ${}^{c}C_\omega$, are invariantive for change of co-ordinates but not for change of gauge. The numbers in brackets (144), etc., imply that ${}^{E}C_\omega$ involves, when $n = 4$, $206 (= 144 + 36 + 6 + 20)$, independent scalars, that ${}^{W}C_\omega$ involves $26 (= 6 + 20)$, independent scalars, and so on. [It is explained below that the statement that ${}^3C_\omega$ involves 36 scalars, merely means that ${}^3C_\omega$ is fully known when the 36 scalars of $Y_\alpha\beta$ below are known; and, again, ${}^4C_\omega$ is fully known when the 144 first derivatives of the 36 scalars are known. In a purely algebraic sense ${}^4C_\omega$ cannot involve more than 96 scalars.]

In what now follows Eddington's symbols

$$\Gamma_{ab}{}^{c},\ {}^*B_{abc}{}^{h},\ {}^*G_{ab},\ R_{ab},\ F_{ab},\ K_{ab,\,c},\ S_{ab,\,c},\ 2\kappa_{a},$$

are the equivalents of our

$$\mathbf{E}_a,\ {}^{E}C_\omega,\ {}^{E}\Omega,\ \tfrac{1}{2}({}^{E}\Omega+{}^{E}\Omega'),\ \tfrac{1}{2}({}^{E}\Omega-{}^{E}\Omega'),\ \Psi_\beta\gamma,\ \Lambda_\beta\gamma,\ \mathbf{E}_\zeta'\zeta,$$

respectively.

In our interpretations, Eddington's "increment due to parallel displacement" has to be distinguished from our (for Weyl's manifold), "increment due to translation." We use $d_a{}^{PD}$ and $d_a{}^{T}$ for these phrases. As with (5) of § 14 above, we have

$$d_a{}^{PD}V_0\beta\eta\gamma = -V_0\alpha\nabla_\theta\,.\,V_0\beta\eta_\theta\gamma - V_0\mathbf{E}_a\beta\eta\gamma - V_0\beta\eta\mathbf{E}_a\gamma.$$

Prof. Eddington remarks that as this is the difference of two invariants it must be invariant, and therefore must be of the form $2V_0\alpha\Psi_\beta\gamma$ (a more general form than ours, after Weyl, $-V_0\beta\eta\gamma\,.\,V_0\lambda\alpha$ in equation (4) of § 14), where $\Psi_\beta\gamma$ is a covariant vector linear and symmetrical in β and γ (contravariant vector dummies), given by

$$2\Psi_\beta\gamma = -\nabla_\theta V_0\beta\eta_\theta\gamma - \mathbf{E}_\beta'\eta\gamma - \mathbf{E}_\gamma'\eta\beta. \tag{4}$$

$\Psi_\beta\gamma$ can easily be expressed uniquely as the sum of two parts, a contractile part involving n scalars and a non-contractile part involving $\frac{1}{2}(n+2)n(n-1)$ scalars. The contractile part is $-n^{-1}\Psi_\zeta\eta^{-1}\zeta\,.\,V_0\gamma\eta\beta$, and is so named because it is the sole part of $\Psi_\beta\gamma$ which is responsible for the contraction $\Psi_\zeta\eta^{-1}\zeta$, a vector. Thus

$$2\Psi_\beta\gamma = -\lambda V_0\beta\eta\gamma + 2\mathbf{T}_\beta\gamma \tag{5}$$

where λ is a covariant vector (namely, $2n^{-1}\Psi_\zeta\eta^{-1}\zeta$), and $\mathbf{T}_\beta\gamma$ is a covariant vector, linear and symmetrical in β and γ, and satisfying the vector equation

$$\mathbf{T}_\zeta\eta^{-1}\zeta = 0. \tag{6}$$

[Digression.—If $Z_\alpha\beta$ is contravariant (and similar remarks hold when it is covariant), and does not satisfy the equation of symmetry $Z_\alpha\beta = Z_\beta\alpha$, there are three kinds of contraction (derived from the scalar form $V_0\gamma^\times Z_\alpha\beta$). The corresponding contractile parts of $Z_\alpha\beta$ are

$$-\eta^{-1}\zeta_1\,.\,V_0\zeta_2Z_{\zeta_1}\zeta_2\,.\,V_0\alpha\eta\beta,\quad \eta^{-1}Z_\zeta'\zeta\,.\,V_0\alpha\eta\beta,\quad -\eta^{-1}Z_\zeta\eta^{-1}\zeta\,.\,V_0\alpha\eta\beta,$$

the three contractions themselves being

$$-\zeta_1\,.\,V_0\zeta_2Z_{\zeta_1}\zeta_2,\qquad Z_\zeta'\zeta,\qquad Z_\zeta\eta^{-1}\zeta.$$

The first and second of these contractions are equal when $Z_\alpha\beta = Z_\beta\alpha$.]

The λ appearing in equation (5) is the λ we met with in Weyl's theory. If we examine what change occurs in $\Psi_\beta\gamma$, or better in $\eta^{-1}\Psi_\beta\gamma$, due to change of gauge, we at once find that in $\eta^{-1}\Psi_\beta\gamma$ the whole change falls on the λ part, no change at all occurring in $\eta^{-1}\Upsilon_\beta\gamma$; that is $\eta^{-1}\Upsilon_\beta\gamma$ is of zero power.

Calculating from (4) the value of

$$\mathrm{V}_0\alpha\Lambda_\beta\gamma = \mathrm{V}_0\alpha\Psi_\beta\gamma - \mathrm{V}_0\beta\Psi_\gamma\alpha - \mathrm{V}_0\gamma\Psi_\alpha\beta$$

we at once find that

$$\Lambda_\beta\gamma = (\Psi_\beta\gamma - \Psi_\beta'\gamma) - \Psi_\gamma'\beta \tag{7}$$

and

$$\mathbf{E}_\beta\gamma = \mathrm{E}_\beta\gamma + \mathrm{Y}_\beta\gamma \tag{8}$$

where

$$\mathrm{Y}_\beta\gamma = \eta^{-1}(\Upsilon_\beta\gamma - \Upsilon_\beta'\gamma - \Upsilon_\gamma'\beta) \tag{9}$$

and $\mathrm{E}_\beta\gamma$ is exactly as in § 14. On the right of (7) the two bracketed terms furnish the skew part of Λ_β and the third term the self-conjugate part. It will be easily seen that

$$\mathrm{Y}_\zeta'\zeta = -\Upsilon_\zeta\eta^{-1}\zeta = 0. \tag{10}$$

From (9) $\mathrm{Y}_\beta\gamma$ is of zero power. It follows that it is just as legitimate to name $-\mathrm{E}_\beta\gamma$ the increment of γ due to the translation $\beta\,(d_\beta{}^{\mathrm{T}}\gamma)$ as it is to name $-\mathbf{E}_\beta\gamma$ the increment of γ due to the parallel displacement $\beta\,(d_\beta{}^{\mathrm{PD}}\gamma)$. In other words Eddington's manifold is simply Weyl's in which there is supposed to exist a given $\mathrm{Y}_\beta\gamma$ of zero power and satisfying the equation $\mathrm{Y}_\zeta'\zeta = 0$.

With our present notation [see (4) of § 16]

$${}^{\mathrm{E}}\mathrm{C}_\omega\gamma = d_\omega{}^{\mathrm{PD}}\gamma = -\mathbf{E}_{\vartheta.\mathrm{V}_{\iota\omega\nabla_\vartheta}}\gamma + (-\mathbf{E}_\alpha\mathbf{E}_\beta + \mathbf{E}_\beta\mathbf{E}_\alpha)\gamma. \tag{11}$$

Putting $\mathbf{E}_\alpha = \mathrm{E}_\alpha + \mathrm{Y}_\alpha$ we get a part involving E_α, E_β only. This is, of course, our former

$${}^{\mathrm{W}}\mathrm{C}_\omega\gamma = d_\omega{}^{\mathrm{T}}\gamma = -\mathrm{E}_{\vartheta.\mathrm{V}_{\iota\omega\nabla_\vartheta}}\gamma + (-\mathrm{E}_\alpha\mathrm{E}_\beta + \mathrm{E}_\beta\mathrm{E}_\alpha)\gamma. \tag{12}$$

Next we get a second part

$${}^{3}\mathrm{C}_\omega\gamma = (-\mathrm{Y}_\alpha\mathrm{Y}_\beta + \mathrm{Y}_\beta\mathrm{Y}_\alpha)\gamma \tag{13}$$

and a third part which may be written

$${}^{4}\mathrm{C}_\omega\gamma = -(\mathrm{Y}_\alpha\gamma)_\beta + (\mathrm{Y}_\beta\gamma)_\alpha \tag{14}$$

where $(\quad)_\beta$ and $(\quad)_\alpha$ denote absolute increments in the sense developed in § 15 above for Weyl's manifold. (14) follows from the beginning of § 17 from which is deducible

$$(\mathrm{Y}_\alpha\gamma)_\beta = -\mathrm{Y}_{\vartheta\alpha}\gamma\,.\,\mathrm{V}_0\beta\nabla_\vartheta + \mathrm{E}_\beta\mathrm{Y}_\alpha\gamma - \mathrm{Y}_\alpha\mathrm{E}_\beta\gamma - \mathrm{Y}_\gamma\mathrm{E}_\beta\alpha.$$

With $n = 4$ (13) involves the thirty-six scalars of $\mathrm{Y}_1\theta$ and (14) involves their 144 first derivatives (∇). The full meaning of our scheme for ${}^{\mathrm{E}}\mathrm{C}_\omega$, given above, will now be evident.

There are numerous simple quadratic forms available to identify, in Prof. Eddington's manner, with the square of Einstein's interval. Not only may we contract (12) and select its self-conjugate part ${}^0\Omega$, but we may similarly contract each of (13) and (14) in two different ways and this by no means exhausts our available simple forms. Besides (13) and (14) we may similarly use $(\mathrm{Y}_\alpha \mathrm{Y}_\beta + \mathrm{Y}_\beta \mathrm{Y}_\alpha)\gamma$ and $(\mathrm{Y}_\alpha\gamma)_\beta + (\mathrm{Y}_\beta\gamma)_\alpha$ where yet another kind of contraction is available (put $\alpha, \beta = \zeta, \eta^{-1}\zeta$).

Returning to the use we made of (3) above, we see that the skew part of ${}^{\mathrm{E}}\Omega\alpha$ is $-\frac{1}{2}\mathrm{V}_1\alpha\mathrm{V}_2\nabla\mathbf{E}_\zeta{}'\zeta$. Since $\mathrm{Y}_\zeta{}'\zeta = 0$,

$$\mathbf{E}_\zeta{}'\zeta = \mathrm{E}_\zeta{}'\zeta = \nabla \log k - \tfrac{1}{2} n\lambda, \tag{15}$$

by (18) of § 15 above. Hence Prof. Eddington's identification of $\mathrm{V}_2\nabla\mathbf{E}_\zeta{}'\zeta$ in the electromagnetic field is virtually the same as Weyl's identification of $\mathrm{V}_2\nabla\lambda$. Also note that ${}^{\mathrm{E}}\Omega - {}^{\mathrm{W}}\Omega\,(= {}^3\Omega + {}^4\Omega)$ is self-conjugate.

One consequence of Eddington's suggestions is that we may quite likely desire to identify Einstein's squared interval with something very different from our fundamental form $\mathrm{V}_0 d\rho\eta d\rho$. Let Einstein's squared interval be $\mathrm{V}_0 d\rho\theta d\rho$. We may wish to put $\theta = \frac{1}{2}({}^{\mathrm{E}}\Omega + {}^{\mathrm{E}}\Omega')$ or $\theta = {}^0\Omega = \frac{1}{2}({}^{\mathrm{W}}\Omega + {}^{\mathrm{W}}\Omega')$ or $\theta = \frac{1}{2}({}^3\Omega + {}^3\Omega')$, the last making θ depend solely on $\mathrm{Y}_\alpha\beta$. With any such identification we require symbols dependent on θ, as are k, ${}^{\mathrm{W}}\Omega$, etc., on η. We shall, therefore, henceforth use

$$\theta,\ l,\ \mathrm{F}_\beta,\ \Theta_\beta,\ \mathrm{C}_\omega,\ \Omega,\ \varpi,\ \mathrm{G},$$

in the θ manifold (Riemann manifold) in the same way as above

$$\eta,\ k,\ \mathrm{E}_\beta,\ \Gamma_\beta,\ {}^{\mathrm{W}}\mathrm{C}_\omega,\ {}^{\mathrm{W}}\Omega,\ {}^{\mathrm{W}}\varpi,\ {}^{\mathrm{W}}\mathrm{G},$$

have been used in the η manifold (Weyl manifold). The relations between θ, l, F_β, etc. are precisely the same as those between

$$\eta,\ k,\ {}^a\mathrm{E}_\beta,\ {}^a\Gamma_\beta,\ {}^a\mathrm{C}_\omega,\ {}^a\Omega,\ {}^a\varpi,\ {}^a\mathrm{G},$$

(${}^a\mathrm{E}_\beta$, ${}^a\Gamma_\beta$ here replacing our earlier $\bar{\mathrm{E}}_\beta$, $\bar{\Gamma}_\beta$).

If θ is identified in some such manner as just suggested the determinant $|\theta|$ may become zero. $|\theta| = 0$ and $|\theta^{-1}| = 0$ represent loci. Do the loci if really existent, constitute the boundaries of electrons and atomic nuclei? On crossing the boundary do we pass to a region with four real dimensions or to a region of two real and two imaginary dimensions (two-dimensional time and two-dimensional space)? Is it on the other hand meaningless to speak about crossing the boundary, much as to an inhabitant of hyperbolic space it would be meaningless to speak of passing across his space's boundary, the absolute?

On the Characteristics of Cylindrical Journal Lubrication at High Values of the Eccentricity.

By T. E. Stanton, D.Sc., F.R.S.

(Received July 17, 1922.)

(A Report from the National Physical Laboratory to the Lubrication Committee of the Department of Scientific and Industrial Research.)

The object of the investigation described in the present paper was to obtain more definite information than is at present available on the nature of the action between two lubricated surfaces when the value of the intensity of the pressure as defined by load/(area of lubricated surface) is considerably higher than that normally attained in current practice with cylindrical journals. In the bath lubrication experiments of Mr. Beauchamp Tower in 1883,* the load at seizing was about 600 lbs. per square inch of the projected area of the bearing. In recent practice with forced lubrication and special lubricants, it has been found possible to increase the load up to 2000 lbs. per square inch without seizing. In the case of worm gears, however, high efficiencies of power transmission are obtained at values of the pressure between the surfaces which are of the order of those of the limiting compressive stresses of the materials, and it has been suggested† that perfect lubrication, by means of a film of oil completely separating the surfaces, does not obtain in such cases, but that the thickness of the oil between the surfaces is of molecular dimensions, so that the interaction of these molecules with the molecules of the surfaces will determine the nature of the frictional resistance to motion which is experienced.

This condition has been termed Boundary Lubrication by Mr. W. B. Hardy.‡ The evidence in favour of the existence of boundary lubrication in the case of worm gears is that the efficiency of power transmission by such gears is very little affected by the viscosity of the lubricant, but depends appreciably on the constitution of the oil and nature of the solid surfaces. That this evidence is conclusive cannot, in the opinion of the writer, be regarded as certain, since it is conceivable that in these tests the thickness of the film and the area of it which is under pressure should vary in such a way as to produce changes in the resistance of the order of those due to variations of viscosity of the lubricant, and further, that the flow .of the material of the worm wheel, which undoubtedly takes place under the extremely high pressures, may account for the observed effect of the nature of the material.

---

* 'Proc. Inst. Mechanical Engineers,' 1885, p. 59.

† 'Report of Lubricants and Lubrication Inquiry Committee,' 1920, p. 26.

‡ 'Roy. Soc. Proc.,' A, vol. 100, p. 550.

For the purposes of the present investigation, direct observations of the action between the teeth of a worm gear would have been preferable, but as this appeared to be impracticable, recourse was made to cases of cylindrical lubrication. It is realised, of course, that the conditions of oil supply to the teeth of a worm gear are entirely different from those obtaining in cylindrical journal lubrication, since, in the former case, any two elements of surface are only momentarily in contact, and are both covered with lubricant before the next contact, whereas the surface of the bearing is always in action, and is dependent for its oil supply on oil brought up by the unloaded side of the journal.

It was thought, however, that if steady conditions of lubrication could be set up in cases in which the arc of contact of the oil film was small, say of the order of 10° to 20°, information might be obtained bearing on the limiting conditions, as regards thickness of film of lubricant, under which the hydrodynamical theory holds, and possibly on the transition from "perfect" to "boundary" lubrication.*

An attempt has accordingly been made to set up the conditions desired, and it has been found that, by increasing the difference between the radius of the bearing and that of the journal to an extent very much in excess of ordinary practice, it is possible to obtain steady conditions of lubrication with an arc of contact of the oil film as small as 15°, the fluid pressure in the film amounting to as much as 3·5 tons per square inch.

From the observed pressure distribution in the film, it has been found possible to predict the angular position of the point of nearest approach of the journal and bearing, and hence, by a comparatively simple calculation, to obtain from the theoretical expression for the pressure slope, the values of the eccentricity of the bearing and the mean viscosity of the lubricant. In the case of a journal 1 inch in diameter, it has been found that the ordinary equations of motion of a viscous fluid hold for flow between the inclined surfaces at a distance apart of 0·00005 inch or 1·2 μ.

In all the cases of cylindrical bearings under steady lubrication conditions which have been examined, the actual measured distribution of pressure has been found to agree fairly closely with the theoretical distribution, obtained from the integration of the equations of motion of a film of lubricant separating two surfaces of the same dimensions, attitude and eccentricity as the existing bearing and journal, the film having the same viscosity as that of the lubricant actually used.

---

* Mr. Hardy uses the term complete lubrication to denote any hydrodynamical state of lubrication. The reason for the use of the term "perfect" in this paper is that Reynolds had previously used the term "complete" to distinguish the case in which the film of lubricant extends over the whole arc of the bearing from that in which it extends over a portion of it ("Theory of Lubrication," 'Phil. Trans.,' Part I, p. 62 (1886).

It was concluded, therefore, that the conditions were in all cases those of perfect lubrication, and no approximation to the hypothetical ones of boundary lubrication have yet been observed.

DESCRIPTION OF APPARATUS.

For the purpose of the experiments, the apparatus illustrated in fig. 1 was constructed. The journal, which is of tool steel, hardened and ground, is 1 inch in diameter and 3 inches long. The bearing consists of a cylindrical phosphor bronze bush, very carefully bored out and lapped, and its true internal diameter determined in the Metrology Department of the Laboratory. Considerable difficulty was experienced in securing the requisite uniformity in the internal diameter of the bush, which was essential, and it is due to the skill of Mr. A. Jenkinson, of the Mechanical Staff of the Department, that success in this respect was secured. In the case of the 1·02-inch bearing used for the majority of the experiments, it was found that at no point did the radius vary by more than 1 per cent. from the mean value.

The values of the pressure in the oil film were obtained by connecting a small hole in the centre section of the bearing to a Bourdon gauge, carried by the housing of the bearing. In estimating pressures by this method, great care was necessary to ensure that all air was extracted from the pipe connections and the tube of the pressure gauge.

As the bearing was in the form of a complete cylindrical bush, it was not thought necessary to attempt to secure the conditions of bath lubrication adopted in Tower's experiments with a half bush, the oil being fed into the bush from a reservoir communicating with the solid end of the bush under a head of about 4 inches. Further, in order to avoid the complication of a large number of pressure holes and pipe connections, all the pressures were measured from one hole, the angular distribution of pressure being obtained by rotating the bearing in its housing through a known angle. The validity of this method depended, of course, on no wear of the bearing taking place during the running, and evidence of wear was carefully sought during the experiments. It was found that, provided the apparatus was always started up under no load, and the load gradually imposed with

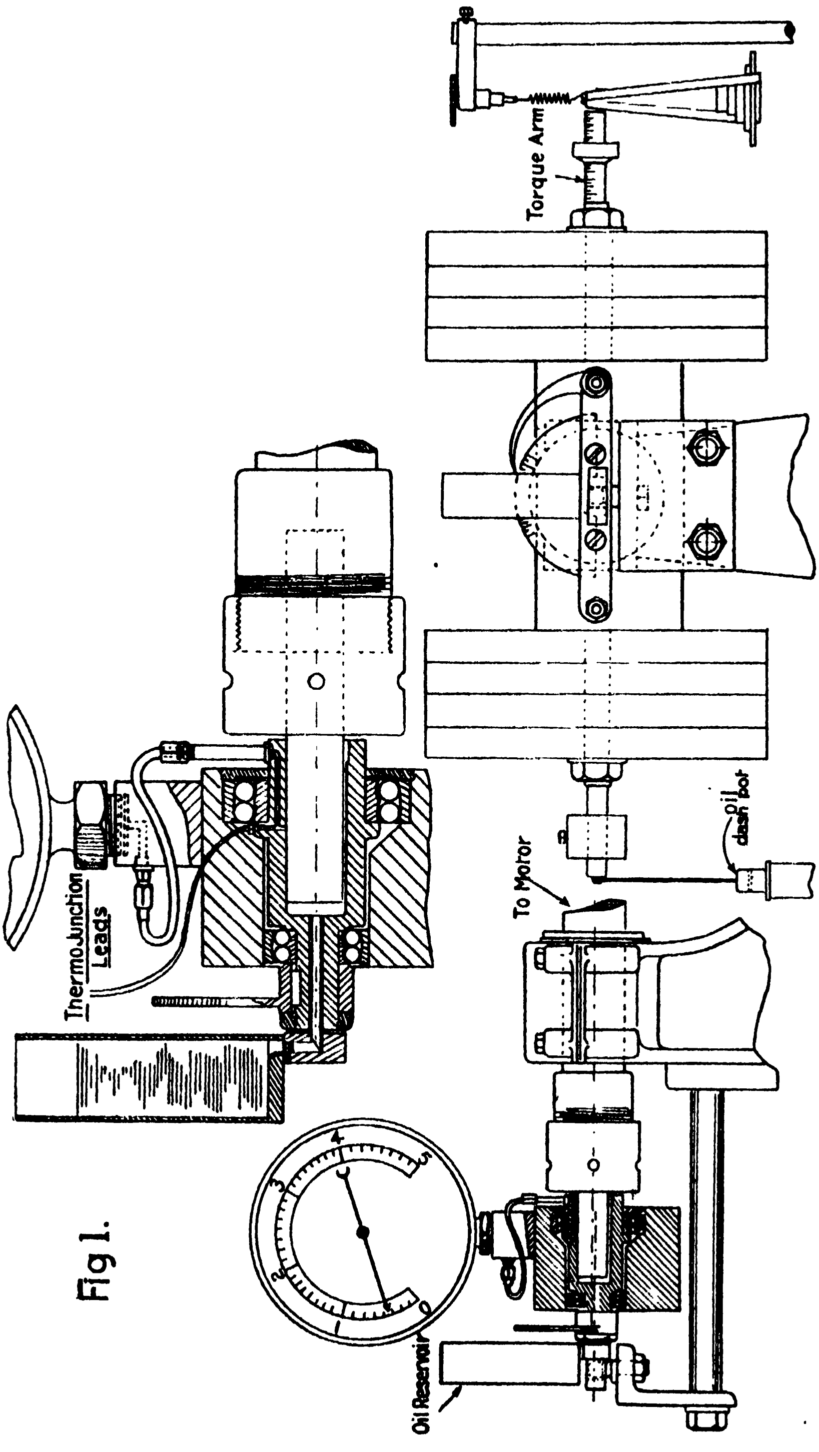

Torque Arm
Thermo Junction Leads
To Motor
oil dash pot
Oil Reservoir
0
1
2
3
4
5

Fig 1.

similar precautions on stopping, no wear could be detected, even after prolonged running. It was therefore assumed that the method of rotating the bearing during a pressure distribution determination was justifiable, and obviously the saving of time due to the method was very great.

In Tower's original experiments no attempt was made to determine the temperature of the oil film, the preliminary assumption being that the value of this was that of the oil bath. It was concluded by Reynolds, in his analysis of Tower's experiments, that this assumption was not justified, and a correction was introduced based on the heat dissipated by the known frictional resistance.

In the present experiments an attempt has been made to measure the temperature of the film by inserting a thermojunction in the pressure hole, after the pressure had been determined, in such a manner that the end of the junction was in contact with the surface of the revolving journal, and a small amount of oil was allowed to leak through the oil-way by means of a special gland. It is thought that by this means the temperature of the oil film has been determined with fair accuracy, but evidence has been found that the temperature of the film is not, as would be expected, uniform throughout.

The frictional torque on the bearing was estimated by direct weighings in the scale-pan of the torque arm attached to the housing of the bearing. Want of accuracy in torque determinations, owing to the effect of static friction on the zero position of the arm was eliminated by running the journal in both directions and taking the mean reading.

RESULTS OF THE EXPERIMENTS.

Case I.—Diameter of journal, 1·00 *inch*; *diameter of bearing,* 1·02 *inch.*

The observed pressure distribution at a speed of 1000 r.p.m. and with a total load of 690 lb. is given in fig. 2 for sperm oil and in fig. 3 for rape oil. The pressure gauge readings are given in Tables I and II. The measured torques with these oils were 0·25 and 0·59 (lb.-inches), giving values of the nominal coefficient of friction of 0·00072 and 0·0017 respectively. It may be remarked that the coefficient for sperm oil is appreciably less than that found in the best ball-bearing practice.

Experiments were also made with more viscous oils, *i.e.*, castor and a mineral oil known as F.F.F. cylinder oil, but in these cases the conditions at the extremities of the oil film were evidently not stable and steady pressure gauge readings were not obtainable, whereas the fluctuations of pressure with the less viscous oils were negligible.

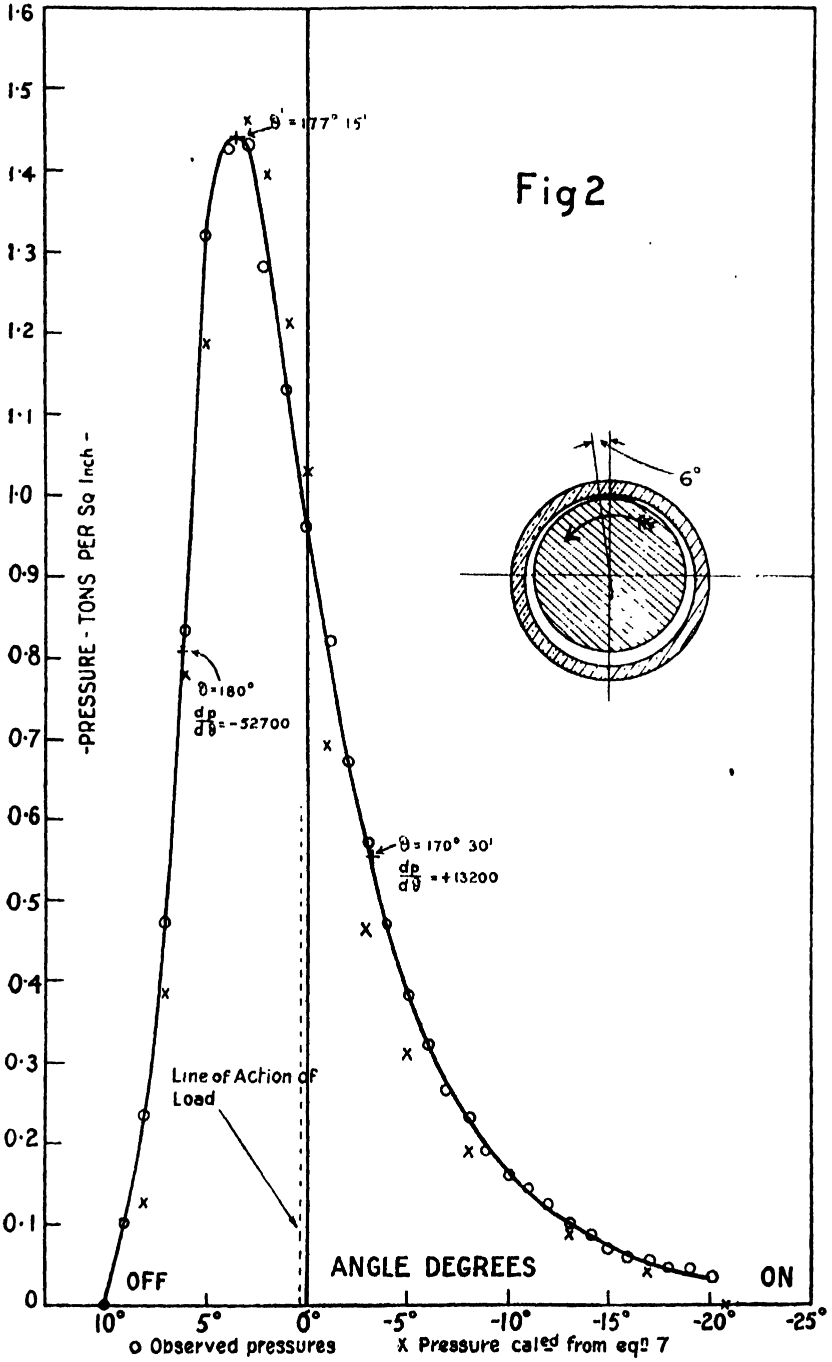

1·6
1·5
1·4
1·3
1·2
1·1
1·0
0·9
0·8
0·7
0·6
0·5
0·4
0·3
0·2
0·1
0
-PRESSURE - TONS PER Sq Inch -
θ' = 177° 15'
θ = 180°
dp/dθ = -52700
θ = 170° 30'
dp/dθ = +13200
6°
Line of Action of Load
OFF
ANGLE DEGREES
ON
10°
5°
0°
-5°
-10°
-15°
-20°
-25°
o Observed pressures
X Pressure calc^ed from eq^n 7

Fig2

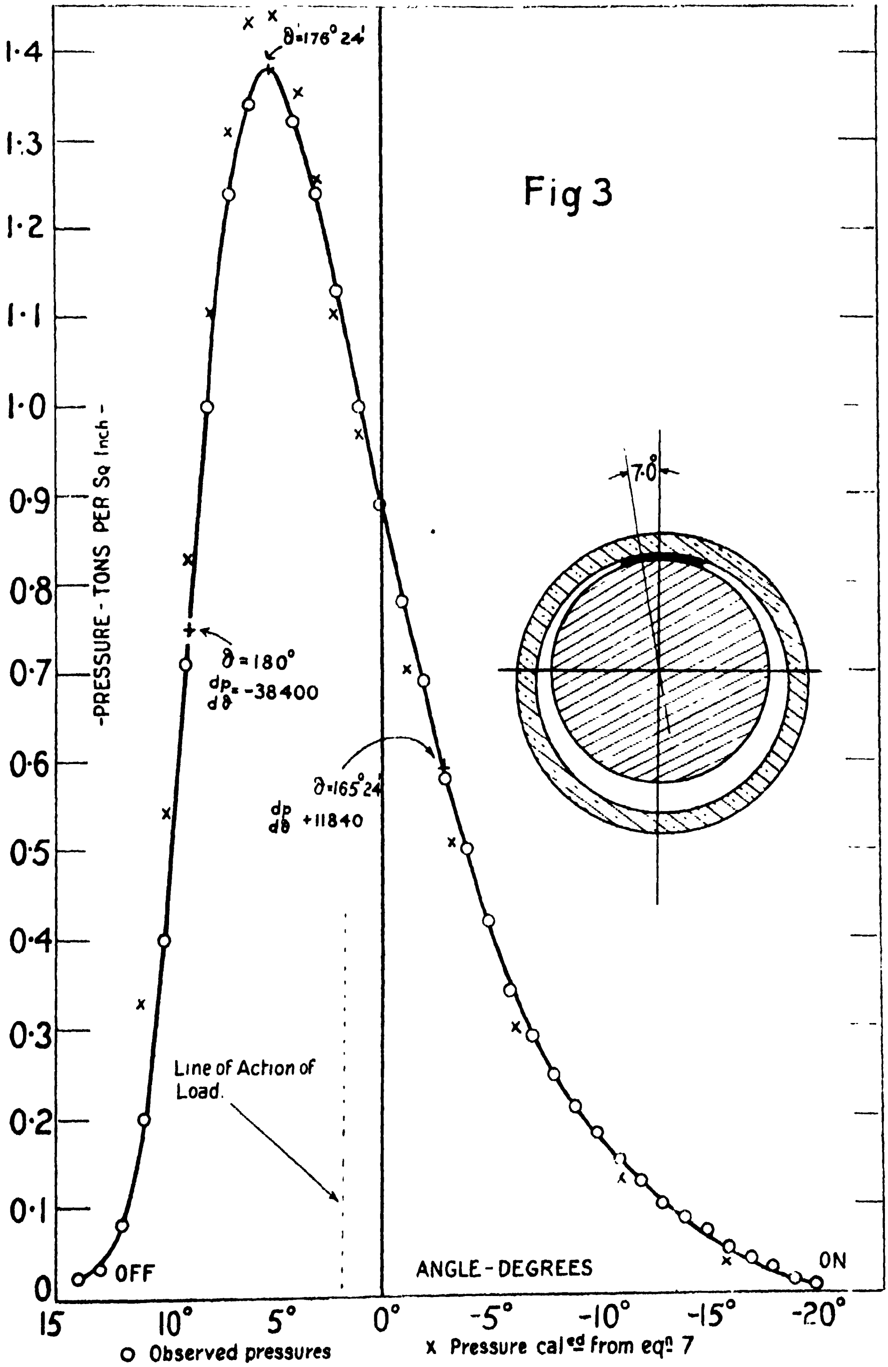


Fig 3
-PRESSURE - TONS PER Sq Inch -
1·4
1·3
1·2
1·1
1·0
0·9
0·8
0·7
0·6
0·5
0·4
0·3
0·2
0·1
0
15
10°
5°
0°
-5°
-10°
-15°
-20°
ANGLE - DEGREES
OFF
ON
θ=176° 24'
θ=180°
dp/dθ = -38400
θ=165° 24'
dp/dθ +11840
Line of Action of Load.
7·0°
O Observed pressures
x Pressure cal^ed from eq^n 7

Case II.—Diameter of journal, 1·00 *inch; diameter of bearing,* 1·06 *inch.*

It was thought that in the case of the more viscous oils steadier pressure conditions would probably be obtained by the use of a still greater difference in the radii of bearing and journal, and a bearing having an internal diameter of 1·06 inch was made. On making a series of observations at the same load and speed as before, it was found that this conclusion was correct, and satisfactory tests were made with castor oil and F.F.F. cylinder oil. The results obtained are shown in figs. 4 and 5 and in Tables III and IV.

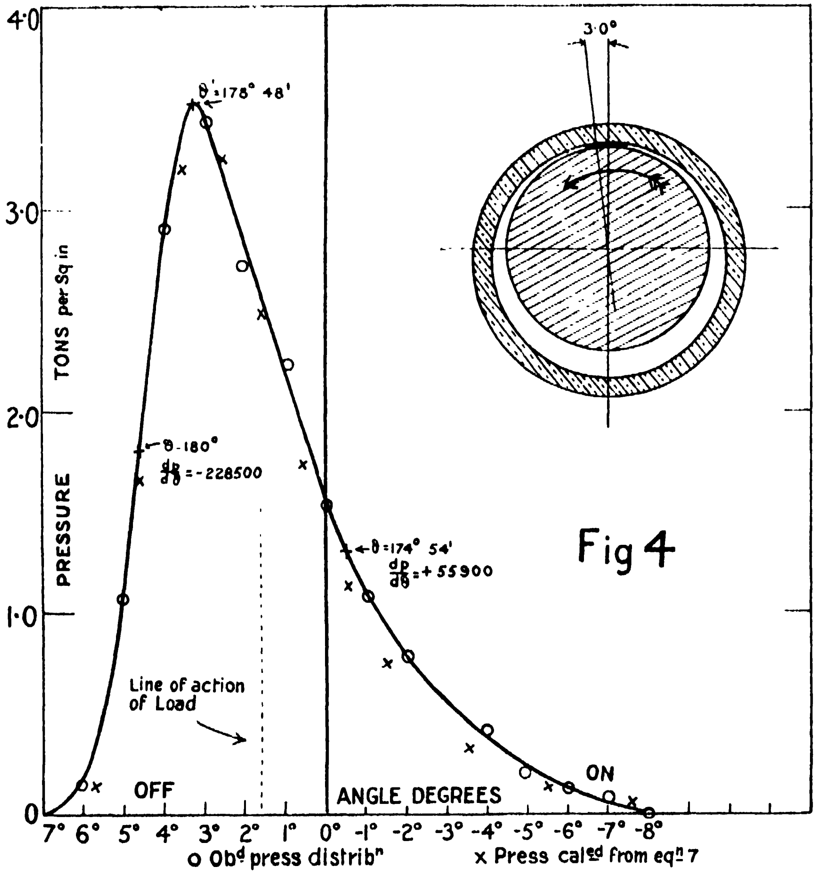


Fig 4

It will be seen that in these experiments steady conditions of lubrication were obtained with an arc of contact of the film of from 14° to 15°, and with fluid pressures up to 3·5 tons per square inch. The nominal coefficients of friction were, however, higher than in the previous cases, the values being 0·0023 for the castor oil and 0·0035 for the F.F.F. cylinder oil.

Experiments were also tried with this bearing at the same load and speed, and using sperm oil as a lubricant, but it was not found possible to attain steady conditions before "seizing" occurred, so that the limit as regards reduction of arc of contact of film with this method of supplying the lubricant appears to have been reached.

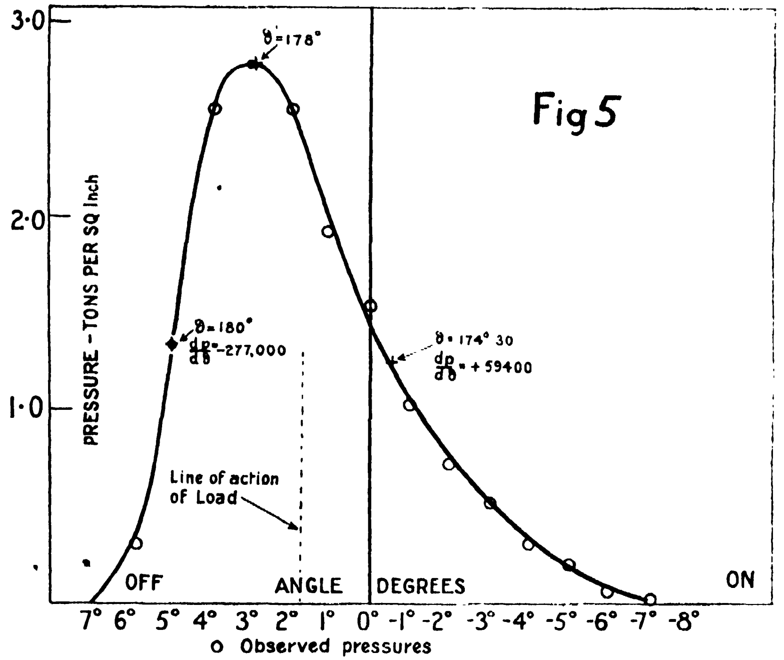


Fig 5

Calculation of the Eccentricity.

In fig. 6 let O be the centre of the journal of radius a, O′ the centre of the bearing of radius $a+\eta$, OA the radius through the point of nearest approach

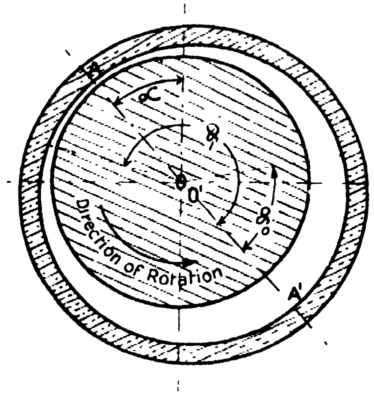


Fig. 6.

making an angle α with the vertical. Then if OO′ $= c\eta$ the distance apart of the surfaces can be written

$$h = \eta\,(1 + c\cos\theta), \tag{1}$$

where θ is measured from OA′ in a counter-clockwise direction. c is called the eccentricity of the bearing.

Assuming the space between the surfaces from θ_0 to θ_1 to be filled with a lubricant in steady motion, it was shown by Reynolds,* from the hydrodynamical theory of a viscous fluid, that the pressure slope in the film is given by

$$\frac{dp}{d\theta} = \frac{6\mu \mathrm{U}ac\,(\cos\theta - \cos\theta')}{\eta^2\,(1 + c\cos\theta)^3}, \tag{2}$$

where θ' is the value of θ at the point of maximum pressure, μ the viscosity of the film, and U the relative velocity of the solid surfaces; and that the intensity of the frictional resistance at the surface of the bearing is

$$f = \left[\frac{\mu \mathrm{U}}{h} - \frac{1}{2}\frac{h}{a}\cdot\frac{dp}{d\theta}\right]_{y=h}. \tag{3}$$

Assuming, therefore, that the bearing is in equilibrium under the load L, acting through O, the fluid pressure and friction, and an externally applied couple, M, per unit of length, we have

$$\int_{\theta_0}^{\theta_1}\{p\sin(\theta+\alpha) + f\cos(\theta+\alpha)\}\,d\theta = 0, \tag{4}$$

$$\int_{\theta_0}^{\theta_1}\{-p\cos(\theta+\alpha) + f\sin(\theta+\alpha)\}\,a\,d\theta = \mathrm{L}, \tag{5}$$

$$\int_{\theta_0}^{\theta_1} f a^2\,d\theta = \mathrm{M}, \tag{6}$$

and another equation representing the boundary conditions of the fluid.

It would appear, therefore, that, provided L, M, U, a, and μ are known, the values of η, c, α, and θ' and the distribution of pressure round the bearing could be determined.

As has been remarked by Harrison,† equation (2) is directly integrable, the solution being

$$\begin{aligned} p = c + \frac{6\mu \mathrm{U}a}{(1-c^2)^2\eta^2}\Bigg[&\frac{c\sin\theta}{(1+c\cos\theta)^2}\{-(1-c^2)(1+c\cos\theta) \\ &+ \tfrac{1}{2}(1+c\cos\theta')\,[1-c^2+3\,(1+c\cos\theta)]\} \\ &+ (1-c^2)^{-\frac{1}{2}}\{2(1-c^2)-(1+c\cos\theta')(2+c^2)\}\tan^{-1}\left\{\left(\frac{1-c}{1+c}\right)^{\frac{1}{2}}\tan\frac{\theta}{2}\right\}\Bigg]. \end{aligned} \tag{7}$$

---

* 'Phil. Trans.,' Part I, p. 194 (1886).

† 'Trans. Cambridge Phil. Soc.,' vol. 22, No. III.

It will be apparent, therefore, that the integration of equations (4), (5), and (6) in a manner suitable for direct evaluation of η, c, α, and θ' would involve extremely laborious and complex calculations, and to avoid this Reynolds resorted to the device of the expansion of $dp/d\theta$ in a series of ascending powers of c, and then integrated term by term. The solution of equations (4), (5), and (6) was thus rendered comparatively simple, but as the series were not convergent for values of c greater than 0·6, the solution thus obtained was not applicable to cases of bearings in which the eccentricity was high. This restriction was not, however, considered by Reynolds to be an important one, for the reason that in the practicable case of a semi-cylindrical bearing he was led to the conclusion that the pressure at the off extremity of the bearing becomes negative when the value of c exceeds 0·5. Since, in his opinion, the existence of negative pressures in the film would probably lead to rupture of the film, this value of c was regarded by Reynolds as the limit of safe load, and therefore, as will be seen from equation (1), "the limiting load for which the half-bearing will run full is that which caused the least distance between the surfaces to be half the difference of the radii." *

For this reason further integration of the equations for cases in which c was greater than 0·5 was not attempted by Reynolds, although he was careful to point out that since "the rupture of the film does not take place at the point of nearest approach, the brass may still be entirely separate from the journal, and could the integrations be effected, it would be as possible to deal with this condition as with that of complete lubrication," * *i.e.*, the half-bearing running full.

As, however, will be clear from the results of the experiments described above, the rupture of the film at the off extremity of the bearing, and the consequent reduction on the arc of contact of the oil film, is by no means the limit of safe load for a bearing supplied with oil in the manner here described, so that the solution obtained by Reynolds is inapplicable to the present cases, in which the value of the eccentricity is of the order of 0·99.

As no direct solution which would enable the values of the eccentricity, attitude, and position of maximum pressure to be calculated from the dimensions and speed of the bearing, the load, frictional moment, and viscosity of the oil, appeared to be available, attention was given to the much simpler problem of utilising the observed pressure distribution curve for the purpose of obtaining the value of the eccentricity, and the subsequent comparison of the values of the calculated pressure at different points of the bearing with those actually observed.

---

* 'Phil. Trans.,' Part I, p. 163 (1886).

The method of calculating the eccentricity finally adopted was as follows:—

From an inspection of the experimental pressure distribution curves, it was noted that the position of the point of inflection of the curve on the off side of the bearing which, from equation (2), will be seen to locate the point of nearest approach, could be predicted with fair accuracy.

Assuming the value of α so obtained, the values of θ and $dp/d\theta$ for the point of nearest approach, and another selected point, usually the point of inflection on the "on" side, were scaled off from the curve. Inserting these values, and that of θ' in the pressure slope, equation (2), two equations were obtained from which the values of c and μ could be calculated.

If the value of μ so determined was found to be in reasonable agreement with that deduced from the measured temperature of the film, the corresponding value of c was taken to be correct.

As an example of the method the calculation of the value of the eccentricity of the 1·02-inch bearing with sperm oil, fig. 2, may be quoted. From the curve the angular positions relative to the arbitrary zero chosen of the point of nearest approach and the maximum pressure were estimated to be $+6^\circ\ 15'$ and $+3^\circ\ 30'$, so that $\theta' = 177^\circ\ 15'$. The value of the pressure slopes at 100° and $170^\circ\ 30'$ were $-52{,}700$ and $+13{,}280$ lbs. per square inch per radian respectively, so that by substitution in (2) we have

$$c = 0\cdot9946.$$

$$\mu = 0\cdot32 \text{ (in C.G.S. units).}$$

At a temperature of 35·8° C., which was the measured temperature of the film, the value of μ for sperm oil at atmospheric pressure is 0·25, so that, assuming a rise of viscosity due to pressure of some 25 per cent., the agreement between the calculated and true values of μ is fair.

The value of $c = 0\cdot9946$ would appear, therefore, to be approximately correct, so that the distance apart of the surfaces at the point of nearest approach, which is $\eta(1-c)$, is, in this case,

$$0\cdot000054 \text{ inch or } 1\cdot37\,\mu\ (\mu = 0\cdot0001 \text{ cm.}).$$

In the case of the 1·06-inch bearing it was found that, owing, presumably, to the high intensity of the oil pressures, a slight warping of the bush took place, with the observed effect that the fluid pressures directly under the main ball bearing (see fig. 1) were appreciably higher than at other sections. The pressures plotted in figs. 4 and 5 refer to this section, and the values of c are calculated from this distribution. The film temperatures were, however, measured at the central section as before, and are, for this reason, somewhat lower in value than the temperatures at the section of maximum pressure. It is considered that this accounts for the fact that the values of μ calculated from the pressure distribution are lower than those corresponding to the observed temperatures, instead of higher, as in the case of the 1·02-inch bearing.

In the case of the 1·06-inch bearing with castor oil the value of c was found to be 0·99845, but, owing to the greater diameter of the bearing, the distance apart of the surfaces, at the point of nearest approach, was not much less than in the smaller bush, its value being

$$0 \cdot 0000465'' \text{ or } 1 \cdot 18\,\mu.$$

CALCULATION OF PRESSURE DISTRIBUTION.

From the values of c and μ obtained in the manner described above, the theoretical changes of pressure from one extremity of the film to the other may be calculated from equation (7). This has been done for the cases illustrated in figs. 2, 3, and 4, the values of the calculated pressure differences being indicated by crosses.

No marked disagreement between the observed pressure distribution and that calculated in the manner described above could exist, since the values of the slope at the maximum and at the two singular points, in each case, are identical, but it is of interest to note that the total rise of pressure in the bearing can be calculated with fair accuracy, and the comparatively slow rate of rise of pressure from the extremity of the "on" side which is found from equation (7), agrees closely with what is observed in practice.

The general results of the present investigation, together with a typical case of Reynold's analysis of Tower's experiments, may be conveniently exhibited in the following tabular manner:—

Experiment.	Diameter (inches).		Lubricant.	Load in lbs. per inch run.	Relative velocity of surfaces (f.p.s.)	Maximum pressure, lbs. per square inch.	Nominal coefficient of friction.	Estimated angular position relative to vertical through centre of journal.		Eccentricity of bearing (calculated).	Least distance apart of surfaces (inches).	Coefficient of viscosity of lubricant, C.G.S. units.	
	Journal.	Bearing.						Point of nearest approach.	Point of maximum pressure.			From observed temp. of oil film at atmospheric pressure.	Calculated from pressure distribution curve.
Tower	4·0	4·003 (calculated).	Olive	1452	15·7	625	0·0013	42·0° to left.	7·0° to left.	0·5200	0·000720		
N.P.L.	1·0	1·02	Sperm	277	4·36	3230	0·0007	6·0° to left.	3°–15′ to left.	0·9946	0·000054	0·252	0·319
N.P.L.	1·0	1·02	Rape	277	4·36	3190	0·0017	7·0° to left.	3°–24′ to left.	0·9904	0·000096	0·540	0·755
N.P.L.	1·0	1·06	Castor	277	4·36	7840	0·0023	3·0° to left.	1°–48′ to left.	0·99845	0·000046	1·43	1·14
N.P.L.	1·0	1·06	F.F.F. cylinder	277	4·36	6250	0·0035	3°–15′ to left.	1°–15′ to left.	0·9978	0·000066	2·09	1·87

Table I.—Observations on 1·02″ Bearing with Sperm Oil. Load 690 lbs. R.P.M. 1000. Torque 0·25 lbs./ins.

Temperature of Film, 35·8° C.

ON SIDE.

Angle, degrees	−20	−19	−18	−17	−16	−15	−14	−13	−12	−11	−10	−9	−8	−7	−6	−5	−4	−3	−2	−1	0
Pressure, tons/sq. in.	0·033	0·04	0·045	0·052	0·06	0·07	0·085	0·10	0·12	0·14	0·16	0·19	0·23	0·265	0·32	0·38	0·47	0·57	0·67	0·82	0·96

OFF SIDE.

Angle, degrees	0	+1	+2	+3	+4	+5	+6	+7	+8	+9	+10
Pressure, tons/sq. in.	0·96	1·13	1·28	1·43	1·43	1·32	0·83	0·47	0·23	0·10	0·0

Table II.—Observations on 1·02″ Bearing with Rape Oil. Load 690 lbs. R.P.M. 1000. Torque 0·59 lbs./ins.

Temperature of Film 33·4° C.

ON SIDE.

Angle, degrees	−20	−19	−18	−17	−16	−15	−14	−13	−12	−11	−10	−9	−8	−7	−6	−5	−4	−3	−2	−1	0
Pressure, tons/sq. in.	0·01	0·015	0·03	0·04	0·05	0·07	0·085	0·10	0·125	0·15	0·18	0·21	0·245	0·29	0·34	0·42	0·50	0·58	0·69	0·78	0·89

OFF SIDE.

Angle, degrees	0	+1	+2	+3	+4	+5	+6	+7	+8	+9	+10	+11	+12	+13	+14	+15
Pressure, tons/sq. in.	0·89	1·00	1·13	1·24	1·32	1·38	1·34	1·24	1·00	0·71	0·40	0·20	0·08	0·03	0·02	0·0

Table III.—Observations on Bearing 1·06″ diam.* with Castor Oil. Load 690 lbs. R.P.M. 1000. Torque 0·8 lbs./ins

Temperature of Film, 49·4° C.†

	ON SIDE.									OFF SIDE.						
Angle, degrees	−8	−7	−6	−5	−4	−3	−2	−1	0	+1	+2	+3	+4	+5	+6	+7
Pressure, tons/sq. in.	0	0·07	0·13	0·22	0·40	0·55	0·78	1·07	1·52	2·23	2·72	3·42	2·90	1·07	0·15	0·0

Table IV.—Observations on Bearing 1·06″ diam.* with F.F.F. Cylinder Oil. Load 690 lbs. R.P.M. 1000.

Torque 1·2 lbs./ins. Temperature of Film, 51·6° C.†

	ON SIDE.								OFF SIDE.						
Angle, degrees	−7	−6	−5	−4	−3	−2	−1	0	+1	+2	+3	+4	+5	+6	+7
Pressure, tons/sq. in.	0	0·07	0·19	0·30	0·51	0·71	1·02	1·50	1·92	2·55	2·78	2·55	1·33	0·32	0·0

* Taken at section of bearing directly under main ball bearing.

† Taken at central section of bearing.

On the Fluorescence of Æsculin.

By J. C. McLennan, F.R.S., Professor of Physics, and Miss F. M. Cale, M.A., University of Toronto.

(Received June 26,—Revised July 11, 1922.)

I. *Introduction.*

In a paper published in 1918, describing observations on thin films of fluorescent solutions, Perrin* presented an extremely interesting explanation of fluorescence. According to him, organic substances when fluorescing are in the process of being destroyed. This destruction he considers to be the cause of the fluorescence. The molecule of the fluorescent substance is supposed not to possess a permanent ability to emit light under stimulation, but to have the power to give out light at the moment of its transformation. On being transformed it is rendered incapable of further fluorescence. Since the molecules of all the fluorescent substances studied by him contained one or more benzene rings, he made the suggestion that fluorescence in the cases investigated was probably due to the rupture of these rings.

R. W. Wood† has also published recently an account of an investigation on the destruction of the fluorescent powers of certain solutions by exposure to sunlight. He found that the fluorescent solutions studied by him could be transformed into coloured non-fluorescent liquids that gave absorption spectra decidedly different from the original solutions. Prolonged exposure to sunlight, moreover, rendered his solutions colourless. In particular, he found that solutions of rhodamine could be bleached without the emission of fluorescent light by maintaining them at a temperature of 100° C., a result which would indicate that the transformation of the molecules of fluorescent solutions may not be as intimately connected with the phenomenon of fluorescence as Perrin has supposed.

In order to add, if possible, to our knowledge of the subject, some experiments were recently made by the writers on the fluorescence of æsculin, and in the course of these experiments it was found that solutions of æsculin could be rendered non-fluorescent not only by exposure to ultra-violet light but also by passing through them in darkness ozonised oxygen. The transformed solutions, too, were found to possess an absorption spectrum different from that of the original solutions.

---

* Perrin, 'Ann. d. Physique,' (9), vol. 10, pp. 133–159, Sept.-Oct. (1918).

† R. W. Wood, 'Phil. Mag.,' vol. 43, pp. 757–765 (1922).

It is known that when the fluorescent power of æsculin is destroyed by the addition of a little acid the fluorescent power of such solutions can be restored by a few drops of an alkali, such as potassium hydroxide. In the case of the solutions of æsculin transformed in our experiments by ultra-violet light or by ozone, it was found that the fluorescent power could not be restored by the addition of the hydroxide. It was also found that solutions of æsculin freed from air and sealed up in evacuated glass tubes could be made to fluoresce for long periods by exposure to ultra-violet light without losing to any measurable extent their power to fluoresce.

When transforming solutions of æsculin, by passing through them in darkness ozonised oxygen, it was found that the transformation was accompanied by the emission of an extremely faint bluish white light.

II. *Apparatus and Method.*

In transforming aqueous solutions of æsculin ($C_{15}H_{16}O_9$) they were exposed in test-tubes of glass or quartz to the light from a powerful quartz mercury-arc lamp. At regular intervals during the exposure the solutions were removed and their fluorescent powers compared with that of the original unexposed stock solution. The quartz mercury-arc lamp used was 75 cm. in length and carried about 10 ampères. Although the light emitted by the lamp was very intense, the solution suffered only a slight rise of temperature during exposure, the distance between the lamp and the test-tube being generally about 15 cm. or 20 cm.

A Nutting spectrophotometer was used to measure the intensity of fluorescence, and the method of using it is shown in fig. 1. The light from a self-regulating carbon-arc lamp (A) was reflected downwards by a mirror (D) into two cells (H) containing the fluorescent solutions to be compared. The fluorescent light from those cells passed through stops (F) and entered the

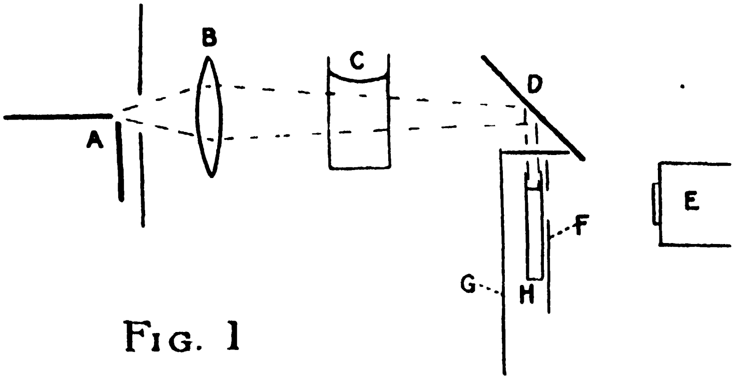


FIG. 1

openings of the photometer in a direction at right angles to the final direction of the exciting beam. By varying the height of the stops the fluorescence at any depth of the liquid could be measured. In our experiments the stops were 4 mm. in width, and the fluorescent powers of the first 4 mm. of the solutions were compared. Solutions sufficiently dilute to show no appreciable absorption over the area investigated were used, so that images uniformly luminous were seen in the photometer. A screen at G served as a dark background for the fluorescent light and kept any stray light emitted by the carbon-arc lamp from entering the photometer directly. The exciting light from the carbon arc was passed through a nickel oxide glass filter, and as a result was limited to the region $\lambda = 3342$ A to $\lambda = 4078$ A with a faint radiation in the extreme red.

In a test made with pure distilled water in the cells used for holding the fluorescent solutions, visible light could only just be detected in the photometer, and that only when the eye was rested in the dark before making the observations. Light of this character was therefore neglected in making comparisons of the intensities of the fluorescent light of the solutions. As the intensity* of the fluorescent light was found to vary directly with the intensity of the exciting light, any slight variation in the intensity of the light from the arc was assumed to affect both the standard æsculin solution and the one under test in the same way.

III. *Sensitivity of Photometer.*

In some preliminary experiments made to test the sensitivity of the photometer, the standard solution was taken as 4/100,000, and it was found that the intensity of fluorescence of solutions as dilute as 1/10,000,000 could be detected and measured.

IV. *Fluorescence and Absorption Spectra of Solutions.*

In an experiment to photograph the fluorescence spectrum of æsculin dissolved in water, the light from a quartz mercury-arc lamp was projected into the solution contained in a fused quartz tube, and the fluorescent light emitted from the tube at right angles to the direction of the exciting light was passed through a quartz spectrograph. The fluorescent spectrum consisted of a broad continuous band, extending from about $\lambda = 5461$ A to $\lambda = 4078$ A.

With long exposures to ultra-violet light the æsculin solutions were transformed from a colourless liquid that exhibited fluorescence into one of a pale amber tint that was non-fluorescent. The absorption spectrum of the fluorescing æsculin solutions was characterised by one broad absorption band in the extreme ultra-violet and a second broad one with maximum absorption close to $\lambda = 3400$ A. In the absorption spectrum of the transformed æsculin solutions the broad band with maximum absorption at $\lambda = 3400$ A was absent. This result indicates that the light responsible for the excitation of the fluorescence of æsculin was practically confined within the limits fixed by the transparency of the nickel-oxide glass filter. The result also shows that in the fluorescence of æsculin solutions we have a definite illustration of the applicability of Stokes's law.

---

* R. W. Wood, *loc. cit.*

By testing solutions that were exposed for different periods attempts were made to see if spectra could be obtained with absorption bands differing slightly from those of the spectrum of the stock solution, but no such spectra were obtained. In this regard æsculin seems to differ from some of the substances studied by R. W. Wood, for, in his experiment on absorption spectra, he appears to have found that, in the transformations which he brought about, an intermediary product was obtained, for which the absorption spectrum was slightly different from that of his stock solution, as well as from that of the completely transformed one.

V. *Experiments on Fluorescence excited by Ultra-Violet Light.*

In the first set of experiments observations were made on the fluorescent powers of solutions of æsculin in water of different concentrations, the strongest solution used being that of four parts by weight of æsculin to 10^6 parts of water. The results of these observations are shown graphically in fig. 2. From the curve it will be seen that for very dilute solutions the

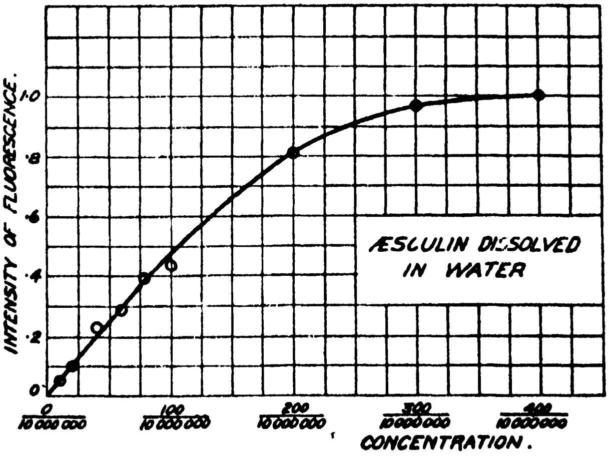


Fig. 2.

fluorescent power was directly proportional to the concentration. For solutions stronger than one part in 10^6 of water this law no longer applied. With solutions having four parts of æsculin to 10^6 parts of water the intensity of the fluorescence reached approximately its maximum value, a result that was in accordance with the observations cited by Perrin.

In a second set of experiments solutions of æsculin in water having the same initial concentrations were placed in tubes of quartz and of glass, having approximately the same thickness of wall, and these were exposed at the same distance from the lighted mercury-arc lamp. The intensity curves for the fluorescence of the solutions in the glass tubes were found to have the same form as those of the solutions in the quartz tubes, but the rate of decay of the fluorescent power of the former was only about one-half of that of the latter. As the glass did not transmit any radiation of wave-length shorter than $\lambda = 3342$ A, it would appear that radiations from the quartz mercury-arc lamp, of wave-lengths shorter than $\lambda = 3342$ A, were at least as effective in transforming the æsculin as the radiations from the lamp having wave-lengths longer than $\lambda = 3342$ A. Illustrations of the curves of decay for the two cases are shown in fig. 3.

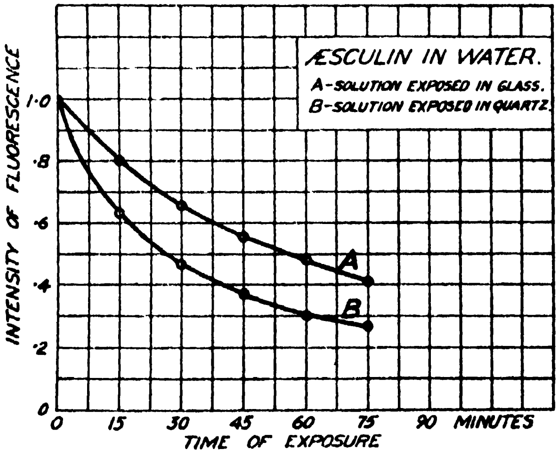


FIG. 3.

In a third set of experiments samples of the stock solution were diluted to the same amount; one by the addition of distilled water, and others by the addition, in varying strengths, of solutions that had completely lost their power of fluorescence through exposure to ultra-violet light. On testing these diluted solutions it was found that they all exhibited the same fluorescent power. From this result the conclusion was drawn that the fluorescent power of untransformed molecules of æsculin in solution in water was not affected by the presence in the solution of molecules of the æsculin that had undergone transformation.

In a fourth set of experiments a solution contained in a glass tube, and having an initial strength of four parts of æsculin in 10^6 parts of water, was exposed for 90 minutes to the ultra-violet light from the powerful quartz mercury lamp referred to above. At intervals during this period the fluorescent power of the solution was measured. At the end of the 90 minutes of exposure the solution was placed in a cabinet in a darkened room for a period of 20 hours, after which the exposure was continued for 45 minutes more, making 135 minutes in all. The solution was then placed once more in the cabinet in the darkened room for 170 hours, and at the end of this time its fluorescent power was again measured. The results of these measurements are shown graphically in fig. 4.

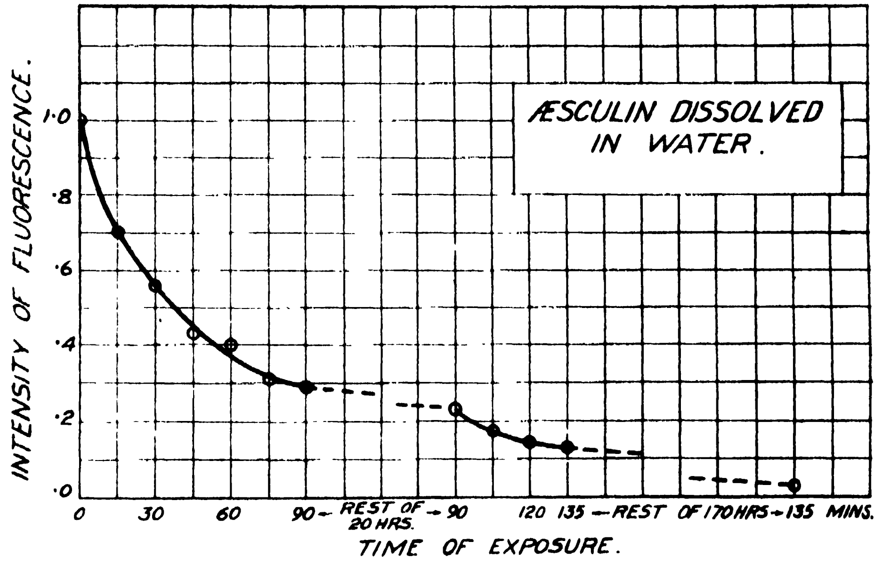


Fig. 4.

By assuming, as a result of the third set of experiments, that the fluorescent power of a solution was not affected by the presence in it of transformed molecules of æsculin, it was possible from the results, to calculate the concentration of untransformed æsculin in the solution at various times during its exposure. The values of these concentrations are shown in fig. 5. From these values the changes in concentration per second were calculated and the percentage changes of the untransformed æsculin in solution were deduced. Curves representing them are given in fig. 6. From these latter curves it will be seen that the percentage rate of transformation decreased rapidly in

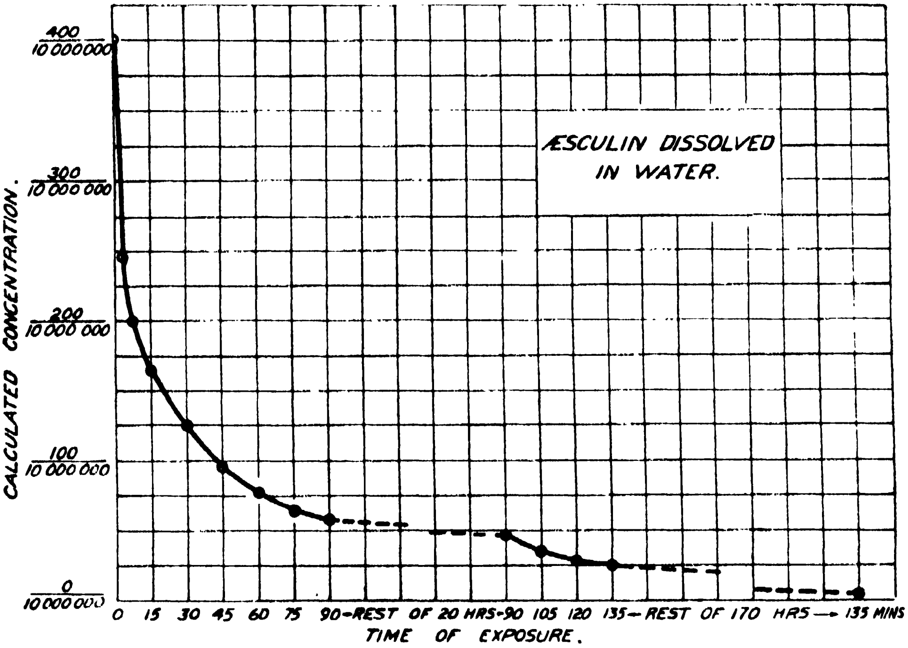

400/10 000 000
300/10 000 000
200/10 000 000
100/10 000 000
0/10 000 000
CALCULATED CONCENTRATION.
ÆSCULIN DISSOLVED IN WATER.
0 15 30 45 60 75 90→REST OF 20 HRS→90 105 120 135→REST OF 170 HRS→135 MINS
TIME OF EXPOSURE.

Fig. 5.

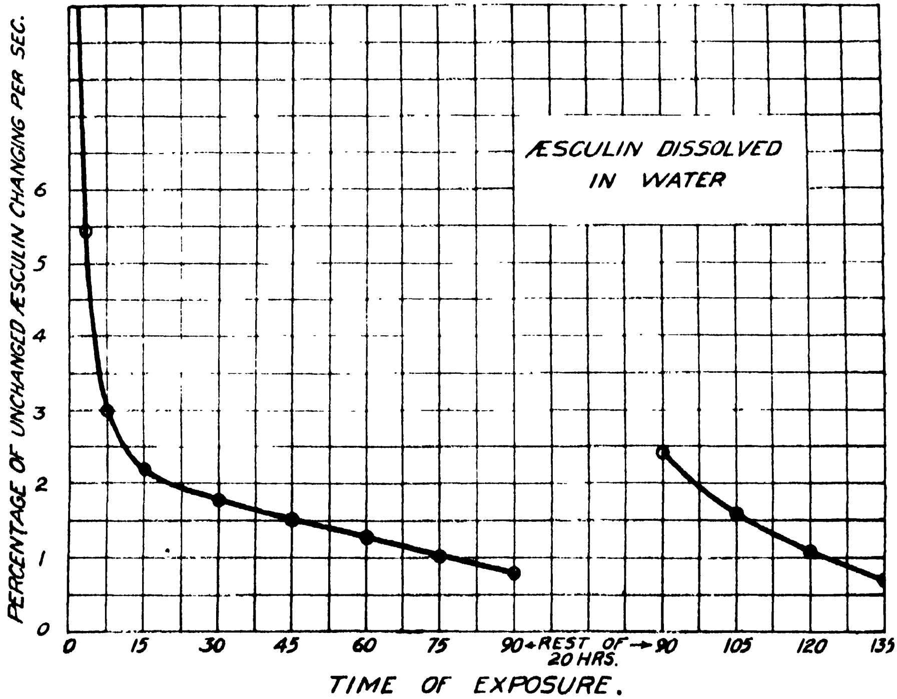

PERCENTAGE OF UNCHANGED ÆSCULIN CHANGING PER SEC.
6
5
4
3
2
1
0
ÆSCULIN DISSOLVED IN WATER
0 15 30 45 60 75 90+REST OF→90 20 HRS. 105 120 135
TIME OF EXPOSURE.

Fig. 6.

the initial stages of the exposure, but more slowly after the solution had been exposed for 15 minutes. For the period of the exposure from 15 to 90 minutes the percentage rate of transformation decreased proportionately with the time of the exposure.

Two points of interest that were revealed by these results are (1) that the solution continued to lose its fluorescent power during the periods of 20 hours and 170 hours that elapsed while it was in the cabinet in the darkened room, and (2) that when the exposure was resumed after the 20 hours' rest in darkness the percentage rate of transformation, as measured by the fluorescent power of the solution, was higher than it was a short time before it was placed in the cabinet at the cessation of the 90 minutes' exposure. This suggested that some product was formed during the exposures that prevented the solution from exhibiting its natural fluorescent power, that also continued to transform the æsculin when exposure ceased, and that finally disappeared during the rest period.

In a fifth set of experiments designed to throw light on this behaviour of aqueous solutions of æsculin exposed to ultra-violet light, it was found that when both dilute and strong solutions were exposed for short periods to ultra-violet light and then placed in the cabinet in the darkened room, the dilute solutions lost their fluorescent power completely in the course of two or three weeks, and the strong solutions in the course of two or three months. In the case of all solutions of æsculin that had lost completely their fluorescent power through exposure to ultra-violet light, it was found that their power of fluorescence was not restored by adding a small quantity of hydroxide to them.

VI. *Experiments with Oxygen and Ozone.*

In the experiments described above the solutions were always in contact with air during their exposure to ultra-violet light, and it was thought that oxygen, or possibly ozone, might have played some part in bringing about the results observed. To test the effect of oxygen on the solutions both air and oxygen were bubbled through them in darkness without the solutions showing any diminution in fluorescent power. Experiments were also made by bubbling both air and oxygen through the solutions in the presence of strong light from the quartz mercury arc lamp. In these experiments control solutions were used through which no gas was bubbled, and all the solutions were in glass vessels during exposure. Tests showed that the loss of fluorescent power was the same for the solutions through which the gases were bubbled as for the control solutions.

An experiment was then made in which oxygen was first passed through an ozoniser and then bubbled through the æsculin solution in darkness. In this case it was found that at the end of 5 minutes during which 270 c.c. of ozonised oxygen were bubbled through the solution the fluorescent power decreased to about 6 per cent. of its original value.

A complete set of readings showing the loss of fluorescent power due to bubbling ozonised oxygen through the æsculin solution were taken and a curve representing them is given in fig. 7. The total duration of the bubbling was only 10 minutes.

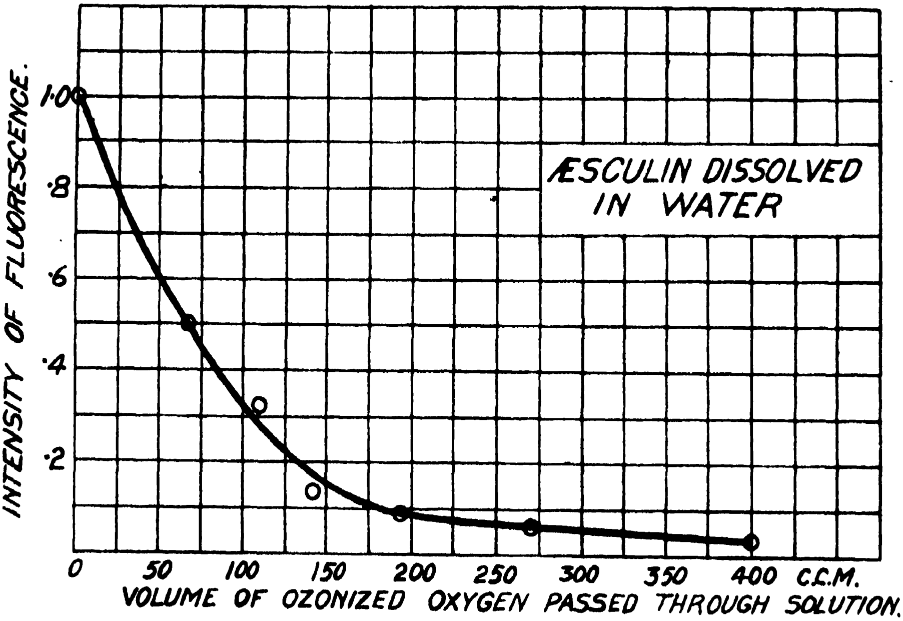


Fig. 7.

When ozonised oxygen was passed through a solution in complete darkness a faint luminescence was emitted by the liquid. With the standard solution, 4/100,000, the luminescence was just visible, but with stronger solutions (*e.g.*, 1/1000) the luminescence was quite distinct and of a bluish-white colour, much like the fluorescence of very dilute solutions of æsculin under the action of ultra-violet light. It was found that when these solutions of æsculin were completely transformed by the action of ozonised oxygen their fluorescent power was not restored by adding a hydroxide to them. In connection with the destruction of fluorescent power by ozone it should be stated that the total emission of light during the process of destruction was infinitesimal compared with the total emission of fluorescent light by solutions of the same initial strength while being transformed by ultra-violet light.

VII. *Experiments with Carbon Dioxide.*

Some experiments were made by bubbling carbon dioxide through solutions of æsculin. In these experiments the fluorescent power of the solutions gradually decreased with the passage through them of the gas. In one particular case, 270 c.c. of carbon dioxide were passed through a solution in 30 minutes and the fluorescent power dropped to 40 per cent. of its initial value. In this case, however, the phenomenon was different from that resulting from exposure to ultra-violet light and from that arising from treatment by ozone for the fluorescent power of the solution was completely restored by adding to it a few drops of alkali solution. The reaction with carbon dioxide was probably a simple acid one.

VIII. *Fluorescence of Æsculin Solutions in Sealed Vessels free of Air.*

(1) *Exposure in Glass Vessels.*—In these experiments the solutions to be exposed were boiled for about a minute to dispel all occluded gases and the glass tubes containing them were exhausted as highly as possible and sealed up. For the purpose of control a solution was boiled for the same time, then poured from one vessel to another to re-dissolve air into it, and finally put into a glass test tube and corked up without being exhausted. The control solution and those in the sealed tubes were placed symmetrically about and at the same distance from the lighted quartz mercury arc lamp. After exposure the seals were broken and the intensity of fluorescence of the solutions measured. A set of results obtained in this way is given in Table I.

Table I.

Solution.	Condition of experiment.	Time of exposure in hours.	Intensity of fluorescence at end of exposure.
A (control)	In corked glass test-tube	3·5	0·32
B	Sealed in evacuated glass tube	3·5	1·00
C	Sealed in evacuated glass tube	30·0	1·00

From these results it will be seen that in neither of the experiments did the solutions sealed in evacuated glass tubes show any indication of loss of fluorescent power through their exposure to the light from the mercury arc lamp. Throughout the whole time of their exposure they continued to fluoresce brilliantly.

(2) *Exposure in Quartz Vessels.*—In these experiments the fluorescent powers of the solutions were first measured and the solutions were then placed

in quartz tubes that were exhausted and sealed up. They were exposed to ultra-violet light in the same way as the solutions in the glass tubes. A set of results is given in Table II.

Table II.

Solution.	Condition of experiment.	Time of exposure (hours).	Intensity of fluorescence at end of exposure.
A	Sealed in evacuated quartz tube	2·0	0·19
B	" " "	3·5	0·10
C	" " "	26·0	0·19
D	" " " (vacuum poor)	4·0	0·02

In these experiments it will be seen that the loss in fluorescent power was considerable. All of the solutions seemed to fluoresce much more brilliantly for the first minute or two of the exposure than they did for the remainder of it. The loss of fluorescent power by solution D was greater than that of the others, but its vacuum was found at the end of the exposure not to be good, and it may be that ozone produced by ultra-violet light from the oxygen of the air in it contributed to the result. The values of the intensity of fluorescence given in the fourth column of Table II for solutions A, C and D were found by breaking the seals of the quartz tubes at the end of the exposures and testing the solutions with the photometer in the manner indicated in the beginning of the paper. In the case of solution B, however, the intensity of fluorescence was measured at intervals while the solution was in the sealed quartz tube, the comparison being made with stock solution in a similar but unsealed and unevacuated quartz tube of the same wall thickness. The complete observations with solution B are recorded graphically in fig. 8. In the same figure a curve is given of the intensities of fluorescence of a control solution contained in an unsealed unexhausted quartz tube.

From the graph, B, it will be seen that for the first hour of exposure the decay of fluorescent power of the sealed solution was small in comparison with that of the unsealed one. The general form of the decay curves was, however, approximately the same in both cases for this period. At the end of the first hour of exposure the decay of the fluorescence of the sealed solution increased, and towards the end of the exposure proceeded approximately linearly with the duration of the exposure.

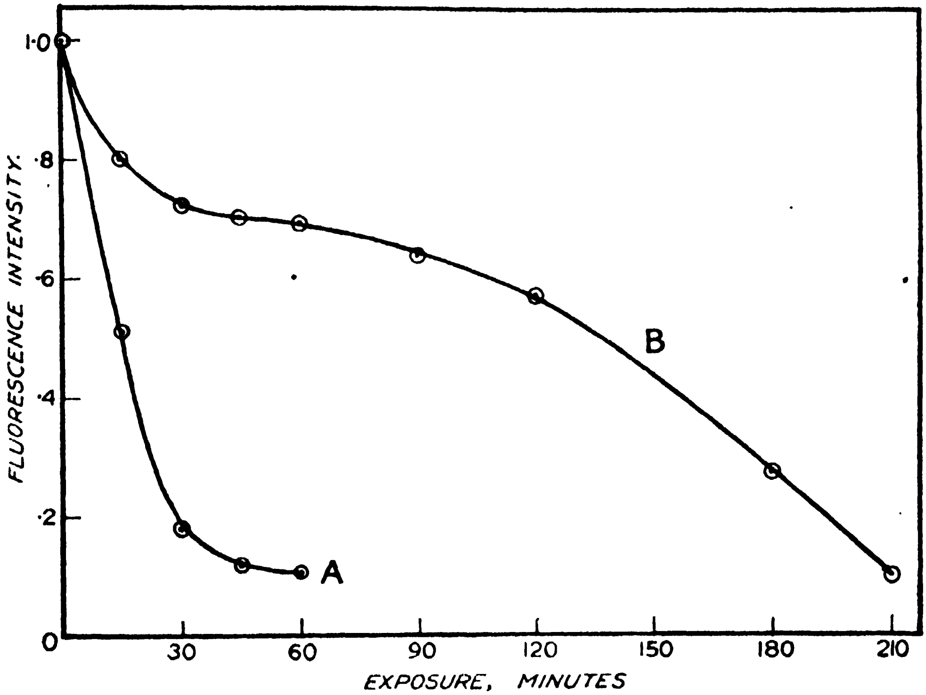


Fig. 8.

IX. *Results of the Investigation.*

The results of the experiments described above show that while the fluorescent power of solutions of æsculin in water can be destroyed by an acid (for example, carbon dioxide) the fluorescent power can be restored by the addition of a few drops of an alkali solution. It has also been shown that the fluorescent power of such solutions can be completely destroyed by passing through them ultra-violet light or ozonised oxygen. In such cases the fluorescent power cannot be restored by the addition of an alkali. The results also show that when ozonised oxygen is slowly bubbled through solutions of æsculin in water, the destruction of fluorescent power is accompanied by a faint luminescence. In this case the phenomenon would appear to be similar to the emission of light during the oxidation of phosphorus. It is also probably similar to the emission of light by decaying wood fibre or decaying organic matter in sea water.

Probably the most interesting results obtained from the investigation are those obtained from observations on the fluorescence of solutions of æsculin in water contained in exhausted sealed tubes. When the solutions were contained in glass tubes they could be made to fluoresce strongly for hours by violet or ultra-violet light in the range of longer wave-lengths, without exhibiting any loss in fluorescent power. When, however, they were exposed in sealed evacuated quartz tubes the fluorescent power was gradually and irrecoverably destroyed by ultra-violet light of the shorter wave-lengths.

The experiments on the fluorescence of aqueous solutions of æsculin contained in evacuated sealed glass tubes are of special importance, for they go to show that molecules of æsculin can be made to fluoresce strongly for hours without being destroyed when subjected to light of suitable wave-length. This result would seem to show that the view put forward by Perrin, namely, that the emission of fluorescent light by an organic substance is evidence of its destruction is not tenable generally.

The Propagation of Gravitational Waves.

By A. S. EDDINGTON, F.R.S.

(Received October 11, 1922.)

1. The problem of the propagation of disturbances of the gravitational field was investigated by Einstein in 1916, and again in 1918.* It has usually been inferred from his discussion that a change in the distribution of matter produces gravitational effects which are propagated with the speed of light; but I think that Einstein really left the question of the speed of propagation rather indefinite. His analysis shows how the co-ordinates must be chosen if it is desired to represent the gravitational potentials as propagated with the speed of light; but there is nothing to indicate that the speed of light appears in the problem, except as the result of this arbitrary choice. So far as I know, the propagation of the absolute physical condition—the altered curvature of space-time—has not hitherto been discussed.

Weyl† has classified plane gravitational waves into three types, viz.: (1) longitudinal-longitudinal; (2) longitudinal-transverse; (3) transverse-transverse. The present investigation leads to the conclusion that transverse-transverse waves are propagated with the speed of light *in all systems of co-ordinates*. Waves of the first and second types have no fixed velocity—a result which rouses suspicion as to their objective existence. Einstein had also become suspicious of these waves (in so far as they occur in his special co-ordinate-system) for another reason, because he found that they convey no energy.

---

* 'Berlin Sitzungsberichte,' p. 688 (1916); p. 154 (1918).

† 'Raum, Zeit, Materie,' 4th edition, p. 228; English edition, p. 252.

They are not objective, and (like absolute velocity) are not detectable by any conceivable experiment. They are merely sinuosities in the co-ordinate-system, and the only speed of propagation relevant to them is "the speed of thought."

The purpose of Einstein's special choice of co-ordinates now becomes clear; they are chosen so as to compel these spurious waves to travel with the same speed as the genuine waves. He imposes conditions on the co-ordinates which bar out most of the spurious disturbances possible; but those which have the velocity of light take advantage of their close resemblance to the genuine waves and slip through the barrier. It is evidently a great convenience in analysis to have all waves, both genuine and spurious, travelling with one velocity; but it is liable to obscure physical ideas by mixing them up so completely. The chief new point in the present discussion is that when unrestricted co-ordinates are allowed the genuine waves continue to travel with the velocity of light and the spurious waves cease to have any fixed velocity.

Plane waves are a very special and artificial case of gravitational propagation, and in the second part of the paper I consider divergent waves. Although the equations of the theory are the same as those occurring in the propagation of sound waves, there is no propagation of gravitation uniformly in all directions like a spherical sound wave. To show the type of divergent waves which can exist, we develop in detail the solution for the waves produced by a spinning rod. We confirm Einstein's result that the rod will slowly lose energy by these waves. Spurious divergent waves can exist just as spurious plane waves can exist; but there is no longer a simple criterion for distinguishing them. The only case in which Weyl's three types of waves exist independently of one another is that of plane waves.

Our calculations show that there is one type of transverse-transverse waves which is inconsistent with the equations $G_{\mu\nu} = 0$. We have examined this case and find that such waves can actually exist, but must be accompanied by other physical manifestations not of a purely gravitational character. This one remaining mode of propagation of gravitational influence turns out to be propagation by light waves.

Plane Waves.

2. Take approximately Galilean co-ordinates and set

$$g_{\mu\nu} = \delta_{\mu\nu} + h_{\mu\nu},$$

where $\delta_{\mu\nu}$ denotes the Galilean values and $h_{\mu\nu}$ the small deviations representing the passage of gravitational waves.

We consider plane waves proceeding with velocity V in the negative direction of the axis of x, so that the $h_{\mu\nu}$ are periodic functions of $(x+Vt)$ only. Writing as usual (x_1, x_2, x_3, x_4) for (x, y, z, t), the argument is (x_1+Vx_4). Denoting differentiation with respect to this argument by an accent, we have

$$\frac{\partial^2 g_{\mu\nu}}{\partial x_1^2} = h_{\mu\nu}'', \qquad \frac{\partial^2 g_{\mu\nu}}{\partial x_1 \partial x_4} = V h_{\mu\nu}'', \qquad \frac{\partial^2 g_{\mu\nu}}{\partial x_4^2} = V^2 h_{\mu\nu}'', \tag{1}$$

and all other second derivatives are zero.

We consider waves of small amplitude, so that the square of the amplitude is neglected. The Riemann-Christoffel tensor is calculated from the formula

$$(\mu\rho\nu\sigma) = B_{\mu\nu\sigma\rho} = \frac{1}{2}\left(\frac{\partial^2 g_{\mu\nu}}{\partial x_\sigma \partial x_\rho} + \frac{\partial^2 g_{\sigma\rho}}{\partial x_\mu \partial x_\nu} - \frac{\partial^2 g_{\mu\sigma}}{\partial x_\nu \partial x_\rho} - \frac{\partial^2 g_{\nu\rho}}{\partial x_\mu \partial x_\sigma}\right), \tag{2}$$

where the terms of the second order of small quantities (products of Christoffel symbols) have been omitted. The 21 different components are easily found from (1) and (2) as follows :—

$$\left.\begin{array}{lll}
(1212) = \frac{1}{2}h_{22}'' & (1224) = -\frac{1}{2}Vh_{22}'' & (2424) = \frac{1}{2}V^2h_{22}'' \\
(1213) = \frac{1}{2}h_{23}'' & (1234) = -\frac{1}{2}Vh_{23}'' = (1324) & (2434) = \frac{1}{2}V^2h_{23}'' \\
(1313) = \frac{1}{2}h_{33}'' & (1334) = -\frac{1}{2}Vh_{33}'' & (3434) = \frac{1}{2}V^2h_{33}'' \\
\multicolumn{3}{c}{(1223), (1323), (1423), (2323), (2324), (2334) = 0} \\
\multicolumn{3}{c}{(1424) = \frac{1}{2}V^2h_{12}'' - \frac{1}{2}Vh_{24}'' \qquad (1434) = \frac{1}{2}V^2h_{13}'' - \frac{1}{2}Vh_{34}''} \\
\multicolumn{3}{c}{(1214) = \frac{1}{2}h_{24}'' - \frac{1}{2}Vh_{12}'' \qquad (1314) = \frac{1}{2}h_{34}'' - \frac{1}{2}Vh_{13}''} \\
\multicolumn{3}{c}{(1414) = \frac{1}{2}V^2h_{11}'' - Vh_{14}'' + \frac{1}{2}h_{44}''}
\end{array}\right\} . \tag{3}$$

Einstein's tensor $G_{\mu\nu}$ is given by

$$G_{\mu\nu} = g^{\sigma\rho}B_{\mu\nu\sigma\rho} = g^{\sigma\rho}(\mu\rho\nu\sigma) = \delta^{\sigma\rho}(\mu\rho\nu\sigma)$$

to the first order of small quantities. For example,

$$\begin{aligned} G_{23} &= \delta^{\sigma\rho}(2\rho 3\sigma) \\ &= -(2131)-(2232)-(2333)+(2434) \\ &= -(1213)-0-0+(2434) \end{aligned}$$

by the antisymmetrical properties of the tensor. Treating the other components similarly, Einstein's equations are :

$$\left.\begin{array}{ll}
G_{11} = -(1212)-(1313)+(1414) = 0 & G_{13} = (1223)+(1434) = 0 \\
G_{22} = -(1212)-(2323)+(2424) = 0 & G_{14} = (1224)+(1334) = 0 \\
G_{33} = -(1313)-(2323)+(3434) = 0 & G_{23} = -(1213)+(2434) = 0 \\
G_{44} = -(1414)-(2424)-(3434) = 0 & G_{24} = -(1214)+(2334) = 0 \\
G_{12} = -(1323)+(1424) \qquad\quad = 0 & G_{34} = -(1314)-(2324) = 0
\end{array}\right\} . \tag{4}$$

Substituting from (3), these become (omitting the common factor $\frac{1}{2}$)

$$\left.\begin{array}{ll} -(h_{22}''+h_{33}'')+V^2h_{11}''-2Vh_{14}''+h_{44}''=0 & V^2h_{13}''-Vh_{34}''=0 \\ -h_{22}''+V^2h_{22}''=0 & -V(h_{22}''+h_{33}'')=0 \\ -h_{33}''+V^2h_{33}''=0 & -h_{23}''+V^2h_{23}''=0 \\ -(V^2h_{11}''-2Vh_{14}''+h_{44}'')-V^2(h_{22}''+h_{33}'')=0 & -h_{24}''+Vh_{12}''=0 \\ V^2h_{12}''-Vh_{24}''=0 & -h_{34}''+Vh_{13}''=0 \end{array}\right\}. \quad (5)$$

These equations can be integrated, *i.e.*, the double accents can be removed. There are no constants of integration, since the $h_{\mu\nu}$ are periodic functions. The equations then reduce to the following seven conditions:—

$$h_{22}+h_{33}=0. \quad (6a)$$

$$(1-V^2)(h_{22}, h_{33}, h_{23})=0. \quad (6b)$$

$$h_{24}=Vh_{12}\,;\ h_{34}=Vh_{13}. \quad (6c)$$

$$h_{44}-2Vh_{14}+V^2h_{11}=0. \quad (6d)$$

3. We can separate the coefficients of the disturbance into three groups, viz.:

Transverse-transverse	h_{22}, h_{33}, h_{23}.
Longitudinal-transverse	h_{12}, h_{13}, h_{24}, h_{34}.
Longitudinal-longitudinal	h_{11}, h_{14}, h_{44}.

It will be seen that the three groups represent disturbances which are propagated quite independently of one another; for example, the presence or absence of longitudinal-longitudinal waves makes no difference to the conditions (6*c*) which have to be fulfilled by the coefficients of the longitudinal-transverse waves. I shall indicate the three classes of waves by the initials TT, LT, LL respectively.

For TT waves, h_{22}, h_{23}, h_{33} cannot all vanish; hence by (6*b*)

$$1-V^2=0.$$

Accordingly, TT waves are propagated with velocity unity, *i.e.*, the velocity of light.

For LL and LT waves, h_{22}, h_{23}, h_{33} are zero, and there is no independent equation determining V. The value of V found from (6*c*) and (6*d*) depends on the coefficients of the disturbance, and has no tendency to approximate to the velocity of light.

With a view to discriminating between genuine disturbances of space-time and spurious waves introduced by using sinuous co-ordinate systems, we return to the consideration of the Riemann-Christoffel tensor (3). In consequence of the conditions (6*c*) and (6*d*), the last five components given in the Table (3) all vanish. The surviving components contain only h_{22}'', h_{23}'', h_{33}''. Accordingly, the Riemann-Christoffel tensor depends solely on the TT waves.

For LL or LT waves the Riemann-Christoffel tensor vanishes, so that space-time is flat and the supposed disturbance is an analytical fiction. The waves will disappear altogether if a suitable change of co-ordinates is made.

The principal invariant of empty space $B_{\mu\nu\sigma\rho}B^{\mu\nu\sigma\rho}$ being composed entirely of the TT coefficients h_{22}'', h_{23}'', h_{33}'', will be propagated with their velocity, viz., the velocity of light. This is the answer to the main question which we have set before ourselves—whether the absolute gravitational influence, independent of co-ordinate systems, is propagated with this speed. This was perhaps a foregone conclusion. If we have an absolute gravitational disturbance which has entirely detached itself from the material system which originated it, only one thing can happen to it. It cannot stay at rest, since there is no absolute rest. It must travel; and since there exists only one speed which is independent of frames of reference, it has no choice but to accept that speed. Our direct calculation verifies this prediction.

4. Since $h_{22}+h_{33} = 0$, typical* plane TT waves are represented by

$$ds^2 = \text{steady value} + (\mathrm{A}\,dy^2 - \mathrm{A}\,dz^2 + 2\mathrm{H}\,dy\,dz)\sin 2\pi(x+t)/\lambda.$$

By re-orientating the axes of y and z, the term in H can be made to disappear, so that

$$ds^2 = \text{steady value} + (dy^2 - dz^2)\,\mathrm{A}\sin 2\pi(x+t)/\lambda.$$

It is of interest to examine what kind of change in the distribution of matter is needed to originate these waves. Let the source be a thin plate lying between $x = \pm\epsilon$ and sending out waves symmetrically in both directions, so that

$$ds^2 = \text{const.} + (dy^2 - dz^2)\,\mathrm{A}\sin 2\pi(t+x+\epsilon)/\lambda \qquad (x < -\epsilon).$$

$$ds^2 = \text{const.} + (dy^2 - dz^2)\,\mathrm{A}\sin 2\pi(t-x+\epsilon)/\lambda \qquad (x > +\epsilon).$$

These can be connected continuously by taking within the plate

$$ds^2 = \text{const.} + (dy^2 - dz^2)\,\mathrm{A}\sin 2\pi t/\lambda \qquad (-\epsilon < x < +\epsilon).$$

In this region the tensor $G_{\mu\nu}$ will not vanish. Our previous calculations of $G_{\mu\nu}$ in (5) can be adapted by noting that $h_{\mu\nu}$ is now a function of t only, and we have only to retain those terms coming from double differentiation with respect to t; these are recognised by the presence of the coefficient V^2.

---

* The most general type is compounded from two of these systems since H need not have the same phase as A.

Thus :—

$$G_{22} = \tfrac{1}{2}h_{22}'' = -\frac{2\pi^2}{\lambda^2}A\sin\frac{2\pi t}{\lambda},$$

$$G_{33} = \tfrac{1}{2}h_{33}'' = \frac{2\pi^2}{\lambda^2}A\sin\frac{2\pi t}{\lambda},$$

and all other components are zero. We have also

$$G = \delta^{22}G_{22}+\delta^{33}G_{33} = 0.$$

The material energy-tensor is given by

$$-8\pi T_{\mu\nu} = G_{\mu\nu}-\tfrac{1}{2}\,\delta_{\mu\nu}G. \tag{7}$$

Hence

$$T_{22} = -T_{33} = \frac{\pi}{4\lambda^2}A\sin\frac{2\pi t}{\lambda},$$

and the other components vanish.*

The components T_{22} and T_{33} are the stresses p_{yy} and p_{zz} in the plate, so that a suitable source of TT waves is a periodic stress alternating between two directions at right angles in the plane of the plate. An equivalent source would be a circular motion of the plate in its own plane.

It may seem remarkable that the only genuine gravitational waves arise from very secondary disturbances in the distribution of matter. Moreover, when (as usual) the velocity of the molecules or of the plate as a whole is small compared with the velocity of light, the components T_{22} and T_{33} are of the second order in this small quantity. The principal component, T_{44}, and the components of the first order, T_{14}, etc., do not excite waves by their oscillations. This is at least partly attributable to the very artificial circumstances of production of *plane* waves. If it seems to us that a backward and forward motion of the plate (variation of T_{14}) ought to produce a more considerable disturbance than mere molecular interchanges without mass-motion, we must reflect that, even on the Newtonian theory of attractions, the to-and-fro motion of an *infinite* plate leaves the gravitational field undisturbed. A small plate would produce waves. Again, we might expect powerful waves to be produced by the alternate creation and destruction of mass (variations of T_{44}); but nature, having made no provision for the propagation of the corresponding disturbance, thereby automatically prevents the construction of such a source.

The problem of plane waves is too special and artificial to illustrate fully the ordinary propagation of gravitational potential. We therefore proceed to consider divergent waves.

---

* This refers to the periodic part of the energy tensor. There must be a steady distribution superposed corresponding to the matter which is executing the small oscillations.

Divergent Waves.

5. The general problem of gravitational disturbance due to a prescribed distribution and motion of matter can be treated in the following way. The $G_{\mu\nu}$ vanish in empty space, and for an assigned density, momentum, etc., of the matter they are given by equation (7); consequently in this problem the $G_{\mu\nu}$ have prescribed values at every point. We shall denote these prescribed values by $[G_{\mu\nu}]$. Consider a set of values of $h_{\mu\nu}$ (vanishing at infinity) obtained by solving the equations

$$\Box\, h_{\mu\nu} = 2[G_{\mu\nu}], \tag{8}$$

where the operator $\Box$ denotes $\partial^2/\partial t^2 - \nabla^2$. It can be deduced* from (8) that

$$\Box\left(\frac{\partial}{\partial x_\alpha} h_\mu{}^\alpha - \tfrac{1}{2}\frac{\partial h}{\partial x_\mu}\right) = 0, \tag{9}$$

where $h_\mu{}^\alpha = \delta^{\alpha\sigma} h_{\mu\sigma}$ and $h = h_\alpha{}^\alpha$. This equation is shown to hold everywhere, in matter as well as empty space, provided only that the assigned distribution $[G_{\mu\nu}]$ is consistent with the laws of conservation of energy and momentum. The integral of (9) (with $h_{\mu\nu}$ vanishing at infinity) is

$$\frac{\partial}{\partial x_\alpha} h_\mu{}^\alpha - \tfrac{1}{2}\frac{\partial h}{\partial x_\mu} = 0. \tag{10}$$

It can now be shown that (8) and (10) together lead to

$$G_{\mu\nu} = [G_{\mu\nu}]. \tag{11}$$

Accordingly the solution of (8) gives values of $h_{\mu\nu}$ which will satisfy the prescribed conditions, and constitutes a solution of the actual problem. Now (8) is the usual equation of wave propagation with unit velocity from sources of strength $2\,[G_{\mu\nu}]$, and is solved by retarded potentials, viz.,

$$h_{\mu\nu} = \frac{1}{2\pi}\iiint [G_{\mu\nu}]_{t-r} \cdot \frac{d\xi\, d\eta\, d\zeta}{r}, \tag{12}$$

where r is the distance from (ξ, η, ζ) to (x, y, z).

The foregoing is essentially Einstein's treatment rearranged to show the rigour of the argument; and (12) constitutes his solution of the problem. It is correct only to the first order of small quantities. It leads to a definite co-ordinate system. Other solutions having a different analytical appearance could be obtained by performing co-ordinate transformations on the $h_{\mu\nu}$ without infringing the condition that they must be small quantities.

Consider the disturbance due to a single source, that is to say a material system confined to a small region at the origin. Then $G_{\mu\nu} = 0$ except in this region. Let

---

* For the intermediate steps in the analysis, see 'Espace, Temps et Gravitation,' Supplement, §43, or my 'Mathematical Theory of Relativity,' §57.

$$\{G_{\mu\nu}\} = \iiint G_{\mu\nu}\, d\xi\, d\eta\, d\zeta$$

taken over the source. Treating r as constant over the source, the solution becomes

$$h_{\mu\nu} = \frac{1}{2\pi r}\{G_{\mu\nu}\}_{t-r}. \tag{13}$$

Thus, for a periodic source, $\{G_{\mu\nu}\} = A_{\mu\nu}\, e^{ipt}$, the solution is

$$h_{\mu\nu} = \frac{A_{\mu\nu}}{2\pi} \cdot \frac{e^{ip(t-r)}}{r}, \tag{14}$$

in agreement with the usual formula for spherical waves in the theory of sound and of electro-magnetism.

I am not sure that any writer has explicitly stated that gravitation is propagated isotropically by spherical waves of this kind; but I think that many like myself must have regarded it as an inference from Einstein's equation (8) which was two obvious to need mentioning. It was a great surprise to me to find that the solution (14) *does not satisfy the law of gravitation.*

To check the solution we shall calculate the component G_{23} at a point on the axis of x_1. When (as here) the $h_{\mu\nu}$ are functions of r and t only, the non-vanishing second derivatives are

$$\frac{\partial^2}{\partial x_1{}^2},\quad \frac{\partial^2}{\partial x_2{}^2},\quad \frac{\partial^2}{\partial x_3{}^2},\quad \frac{\partial^2}{\partial x_4{}^2},\quad \frac{\partial^2}{\partial x_1 \partial x_4}.$$

The other five vanish on the x_1-axis. Hence by (2) we find

$$(1213) = \tfrac{1}{2}\frac{\partial^2 h_{23}}{\partial x_1{}^2} \qquad (2434) = \tfrac{1}{2}\frac{\partial^2 h_{23}}{\partial x_4{}^2},$$

so that by (4) the condition $G_{23} = 0$ becomes

$$\tfrac{1}{2}\left(\frac{\partial^2 h_{23}}{\partial x_4{}^2} - \frac{\partial^2 h_{23}}{\partial x_1{}^2}\right) = 0,$$

which is the equation for plane waves and is not satisfied by (14) or even by the more general form (13).

We can calculate similarly the conditions $G_{12} = 0$, $G_{24} = 0$, and with a little more trouble isolate the component h_{24}. It appears that it also has to satisfy conditions inconsistent with (13) or (14). It follows by symmetry that the solution also fails for h_{12}, h_{13}, h_{14}, h_{24}.

I have not isolated the components h_{11}, h_{22}, h_{33}, h_{44}, but we shall see later that the periodic solution fails for these also. It seems that the only case in which (13) satisfies the law of gravitation is the static case of an undisturbed particle. Thus although the equations of propagation appear to be exactly analogous to the equations of the theory of sound, there is a remarkable difference in the type of propagation since isotropic spherical waves of gravitation cannot occur.

6. In our discussion of plane waves we did not explicity exclude waves of supernatural origin; it was sufficient that the waves themselves should obey the law of gravitation. But in deducing Einstein's solution in § 5 it was necessary to insert the condition that the sources obeyed the conservation of energy and momentum. There can be no doubt that the general solution (12) is correct, and the cause of its apparent failure must be that we have been applying it to a kind of source which cannot exist consistently with the laws of conservation. There cannot be a *simple* source of gravitational waves.

At first sight this conclusion seems contrary to common experience; we have no difficulty in constructing a simple source of G_{22}, G_{23}, G_{33} by spinning a rod in the yz-plane. But in the spinning rod the simple source for these components is accompanied by a *doublet* source of momentum G_{24}, G_{34}. The whole disturbance is therefore not a simple source, and the complete system of waves is not of the symmetrical type (14).

It can be deduced generally from equation (10) that simple waves of type (14) for any component of $h_{\mu\nu}$ must always be accompanied by doublet waves for some other component.

The reason for the breakdown of the analogy between sound waves and gravitational waves is now clear. Sound waves satisfy $\Box\phi = 0$, unconditionally; gravitational waves satisfy $\Box g_{\mu\nu} = 0$, but not unconditionally. And, owing to this condition, the most common solution for sound waves is not a solution for gravitational waves. The condition is superfluous in practical problems, because it is the condition that the sources shall be such as can occur in practical problems; but it none the less limits the number of admissible solutions.

I had hoped to generalise the results for plane waves by making a corresponding discussion for simple spherical waves; but we now see that simple spherical waves do not exist. I turn, therefore, to the consideration of spherical waves of a more complex type, which are certainly possible since they emanate from a material source which can actually be constructed.

We shall calculate the complete solution for a rod spinning in the plane of yz. Let $2a$ be the length, σ the mass per unit length, and ω the angular velocity. Choose an instant when the length of the rod is along the y-axis. The velocity, being in the z direction, will contribute to the component T_{33} of the energy-tensor. The integrated amount is $\{T_{33}\} = \frac{2}{3}\sigma\omega^2a^3$. There will necessarily be a tension in the spinning rod; this is in the y direction and contributes to T_{22}. The integrated amount is found to be $\{T_{22}\} = -\frac{2}{3}\sigma\omega^2a^3$.

For any other position of the rod the new values can be calculated by the tensor-law of invariance

$$T_{22}dy^2+2T_{23}dy\,dz+T_{33}dz^2 = T_{22}'dy'^2+2T_{23}'dy'\,dz'+T_{33}'dz'^2.$$

Setting $dy = dy'\cos\omega t-dz'\sin\omega t$, $dz = dy'\sin\omega t+dz'\cos\omega t$, the new coefficients when the rod makes an angle ωt with the axis of y are found to be

$$\{T_{22}\} = -\{T_{33}\} = -\tfrac{2}{3}\sigma\omega^2a^3\cos 2\omega t, \quad \{T_{23}\} = -\tfrac{2}{3}\sigma\omega^2a^3\sin 2\omega t.$$

Let

$$\gamma_{\mu\nu} = h_{\mu\nu}-\tfrac{1}{2}\delta_{\mu\nu}h.$$

Then by (7) and (14) we obtain

$$\gamma_{22} = f \qquad \gamma_{23} = -if \qquad \gamma_{33} = i^2f, \tag{15a}$$

where

$$f = \tfrac{2}{3}\sigma p^2a^3\frac{e^{ip(t-r)}}{r}, \qquad p = 2\omega.$$

As already explained these simple waves from the stress components must be accompanied by doublet waves from the momentum components, which in turn require quadruplet waves from the energy-component. We can calculate these waves most easily from equation (10) which is equivalent to $\partial\gamma_\mu^\alpha/\partial x_\alpha = 0$. Hence

$$\frac{\partial\gamma_{24}}{\partial t} = \frac{\partial\gamma_{22}}{\partial y}+\frac{\partial\gamma_{23}}{\partial z}, \qquad \frac{\partial\gamma_{34}}{\partial t} = \frac{\partial\gamma_{23}}{\partial y}+\frac{\partial\gamma_{33}}{\partial z}, \qquad \frac{\partial\gamma_{44}}{\partial t} = \frac{\partial\gamma_{24}}{\partial y}+\frac{\partial\gamma_{34}}{\partial z}.$$

These give by (15a)

$$\left.\begin{aligned}
\gamma_{24} &= \frac{1}{ip}\left(\frac{\partial}{\partial y}-i\frac{\partial}{\partial z}\right)f\\
\gamma_{34} &= -\frac{1}{p}\left(\frac{\partial}{\partial y}-i\frac{\partial}{\partial z}\right)f\\
\gamma_{44} &= -\frac{1}{p^2}\left(\frac{\partial}{\partial y}-i\frac{\partial}{\partial z}\right)^2f\\
\gamma\ &= \gamma_{44} = -h
\end{aligned}\right\} \tag{15b}$$

If we write

$$u = \frac{1}{ip}\left(\frac{\partial}{\partial y}-i\frac{\partial}{\partial z}\right)$$

the complete set of values of $h_{\mu\nu}$ is given by the scheme

$$\begin{array}{rcccc}
h_{\mu\nu} = & \tfrac{1}{2}u^2 & 0 & 0 & 0\\
& 0 & 1+\tfrac{1}{2}u^2 & -i & u\\
& 0 & -i & -1+\tfrac{1}{2}u^2 & -iu\\
& 0 & u & -iu & \tfrac{1}{2}u^2
\end{array}$$

all operating upon f. These values constitute the solution of the problem. They satisfy (to the first order) the law of gravitation in empty space because

(8) and (10) are both satisfied. Moreover, the values of $\gamma_{22}, \gamma_{23}, \gamma_{33}$ correspond to the spinning rod; and since the values of $\gamma_{24}, \gamma_{34}, \gamma_{44}$ depend on these uniquely, they must also correspond—if the existence of a spinning rod is not inconsistent with the laws of nature.

Divergent waves can be created analytically by transforming the co-ordinate-system so that many other expressions for the solution may be given. We can, for example, remove the "momentum" waves h_{24}, h_{34}, but the other waves then become more complex. Make the transformation

$$x_a' = \delta_\mu{}^a x_\mu - \delta^{\mu a}\xi^\mu,$$

where ξ^μ represents arbitrary functions of the co-ordinates of the first order of small quantities. From the transformation law of $g_{\mu\nu}$, we find

$$h_{\mu\nu}' = h_{\mu\nu} + \frac{\partial \xi^\mu}{\partial x_\nu} + \frac{\partial \xi^\nu}{\partial x_\mu}.$$

If

$$\xi^1 = 0, \qquad \xi^2 = \frac{1}{p^2}\left(\frac{\partial}{\partial y} - i\frac{\partial}{\partial x}\right)f, \qquad \xi^3 = -\frac{i}{p^2}\left(\frac{\partial}{\partial y} - i\frac{\partial}{\partial z}\right)f, \qquad \xi^4 = 0,$$

we find that $\gamma_{24}', \gamma_{34}', \gamma_{44}'$ disappear, but quadruplet waves are introduced into all the other components. Another useful co-ordinate-system is obtained by making half the above transformation. This has the great advantage that it makes h vanish, so that $\gamma_{\mu\nu} = h_{\mu\nu}$, and many of the formulæ of the theory are greatly simplified. The volume of any four-dimensional region is now undisturbed by the waves.

The chief point of practical interest in the problem of the spinning rod is the question whether its energy is gradually carried away by the gravitational waves which are created, so that the rotation would slow down and ultimately stop of its own accord. We shall use our solution to examine this question, which has also been treated more generally by Einstein. We agree with his conclusion that the energy will actually be radiated away, at a very slow rate.

The expression for the pseudo-energy-tensor of the gravitational field, correct to the *second* order, is*

$$32\pi t_{\mu\nu} = \left(\frac{\partial \gamma_{\alpha\beta}}{\partial x_\mu}\frac{\partial \gamma^{\alpha\beta}}{\partial x_\nu} - \tfrac{1}{2}\frac{\partial \gamma}{\partial x_\mu}\frac{\partial \gamma}{\partial x_\nu}\right) - \tfrac{1}{2}\delta_{\mu\nu}\left(\frac{\partial \gamma_{\alpha\beta}}{\partial x_\sigma}\frac{\partial \gamma^{\alpha\beta}}{\partial x^\sigma} - \tfrac{1}{2}\frac{\partial \gamma}{\partial x_\sigma}\frac{\partial \gamma}{\partial x^\sigma}\right), \tag{16}$$

where $\gamma^{\alpha\beta} = -\gamma_{\alpha\beta}$ for the momentum components and $+\gamma_{\alpha\beta}$ for the other components.

$x^\sigma = -x_\sigma$ for the space components and $+x_\sigma$ for the time component.

---

* This is Einstein's expression (second paper, p. 158), modified because Einstein employs imaginary time. It applies to his special system of co-ordinates employed in our solution (15*a*, *b*), and must not be used for any other system.

The radial component of the flow of energy per unit area per unit time (Poynting vector) is obtained by resolving $(t_4^1,\ t_4^2,\ t_4^3)$ along the radius. Thus:—

$$t_4^r = t_4^1 \frac{\partial x_1}{\partial r} + t_4^2 \frac{\partial x_2}{\partial r} + t_4^3 \frac{\partial x_3}{\partial r},$$

so that by (16)

$$32\pi t_4^r = -\left(\frac{\partial \gamma_{\alpha\beta}}{\partial t}\frac{\partial \gamma^{\alpha\beta}}{\partial r} - \tfrac{1}{2}\frac{\partial \gamma}{\partial t}\frac{\partial \gamma}{\partial r}\right)$$

which in our problem simplifies to

$$= -\left(\frac{\partial \gamma_{22}}{\partial t}\frac{\partial \gamma_{22}}{\partial r} + 2\frac{\partial \gamma_{23}}{\partial t}\frac{\partial \gamma_{23}}{\partial r} + \frac{\partial \gamma_{33}}{\partial t}\frac{\partial \gamma_{33}}{\partial r} - 2\frac{\partial \gamma_{24}}{\partial t}\frac{\partial \gamma_{24}}{\partial r} - 2\frac{\partial \gamma_{34}}{\partial t}\frac{\partial \gamma_{34}}{\partial r} + \tfrac{1}{2}\frac{\partial \gamma_{44}}{\partial t}\frac{\partial \gamma_{44}}{\partial r}\right). \tag{17}$$

To obtain the total flow across an infinite sphere, we have to evaluate this expression to the order $1/r^2$; each separate factor is therefore only required to the order $1/r$. Accordingly, (15*a*, *b*) becomes

$$\gamma_{22} = \mathrm{A}\cos p(t-r), \qquad \gamma_{23} = \mathrm{A}\sin p(t-r), \qquad \gamma_{33} = -\mathrm{A}\cos p(t-r),$$

$$\gamma_{24} = \mathrm{A}\left(-\frac{y}{r}\cos p(t-r) - \frac{z}{r}\sin p(t-r)\right),$$

$$\gamma_{34} = \mathrm{A}\left(-\frac{y}{r}\sin p(t-r) + \frac{z}{r}\cos p(t-r)\right),$$

$$\gamma_{44} = \mathrm{A}\left(\frac{y^2-z^2}{r^2}\cos p(t-r) + \frac{2yz}{r^2}\sin p(t-r)\right),$$

where $\mathrm{A} = \frac{2}{3}\sigma p^2 a^3/r$, and can be treated as constant in differentiating.

We find

$$32\pi r^2 t_4^r = (\tfrac{2}{3}\sigma p^3 a^3)^2\left[\frac{2x^2}{r^2} + \frac{1}{4}\frac{(y^2+z^2)^2}{r^4} + \frac{y^4+z^4-6y^2z^2}{4r^4}\cos 2p(t-r) + \frac{yz(y^2-z^2)}{r^4}\sin 2p(t-r)\right]. \tag{18}$$

Taking the average values over the whole sphere, the periodic terms are found to vanish, and the total radiation, $4\pi r^2 t_4^r$, is at the steady rate

$$\tfrac{1}{10}(\tfrac{2}{3}\sigma p^3 a^3)^2,$$

or, if I is the moment of inertia of the rod, the result is

$$\tfrac{32}{5}\mathrm{I}^2\omega^6. \tag{19}$$

I find that Einstein's general formula (*loc. cit.*, equation (30)) applied to our problem gives the result $\frac{16}{5}\mathrm{I}^2\omega^6$. The discrepancy is due to a numerical slip in one or other investigation, and is not of much importance.*

---

* I think that the factor $\frac{1}{2}$ was introduced by Einstein in order to eliminate the energy of the imaginary waves contained in the expression e^{ipt}; and it should have been dropped in a formula where imaginary waves are not introduced.

If C.G.S. units are used, we must multiply the expression (19) by $2{\cdot}7{\cdot}10^{-60}$. For example, a rod of mass 1 kgrm., length 2 metres, making 50 revolutions per second, is found to lose $3{\cdot}10^{-35}$ of its energy of rotation in a year.

The radiation depends on the asymmetry of the source; a circular disc spinning in its own plane would not lose any energy. The formula (19) would evidently apply to other linear distributions, such as two balls connected by a string. We might be tempted to apply it to pairs of masses not connected by material tension, *e.g.*, a hydrogen atom or a double star. For a hydrogen atom the rate of decay is much the same as for the rod, *i.e.*, of the order 10^{-35} per year; for a double-star system it is rather larger, about 10^{-30} per year. It is rather curious that the difference between these two systems at the opposite extremes of magnitude-scale should not be greater. But both applications are probably illegitimate—the atom, because it is complicated by quantum conditions outside our analysis; the star, because the tension (being now a gravitational force) is limited to a small quantity of the first order, and the problem is thus carried out of range of our approximation. But it seems likely that the radiation (if any) will not exceed that given by (19).

There is clearly no practical objection to the existence of this small radiation from rotating systems, and I can see no theoretical reason for not admitting it.

According to (17) the momentum components gave an inflow of energy which neutralises a considerable part of the outflow due to other components. Also, by (18), the flow of energy is mainly along the axis of rotation of the rod, being eight times more intense along the axis than at right angles to it. But these results have no physical interpretation. The various components of the potential have no separable existence, and the Poynting vector is not a true vector. As Einstein has pointed out, the investigation of this paragraph gives the total loss of energy of the material system correctly, but the intermediate steps are merely analytical. The lost energy is not localisable anywhere.

We found that for plane waves the genuine waves arise from the stress-condition of the material source, and we expressed surprise that so inconspicuous a feature of material distribution should be the occasion of them. Our study of divergent waves confirms this. Apparently, the only originating sources (as distinct from scattering sources) are typified by the spinning rod or by an oscillator (*e.g.*, two masses connected by a spring). We cannot disturb the world by creating mass nor by transporting mass within the same source, since the laws of mass and momentum render these processes undisturbable. The simplest interference we can make is to transport momentum, and that is the process represented by the stress-components.

Electromagnetic Gravitational Waves.

8. The three coefficients for plane TT waves may be grouped as follows:—

$$h_{23}, \qquad h_{22}-h_{33}, \qquad h_{22}+h_{33}.$$

The propagation of the first two has already been studied; but according to (6*a*) propagation of the third group is impossible.

We shall now show that waves of the third kind can exist in space which, although not empty in the strictest sense, is empty of all matter. They are in fact the ordinary light waves. It is of interest to note that this electromagnetic mode of propagation of gravitational influence fits into the only gap left unfilled in the previous discussion of plane gravitational waves.

Plane polarised electromagnetic waves travelling in the negative direction of x are represented by the electromagnetic potential.

$$\kappa_2 = \mathrm{A}\sin p(t+x) = \mathrm{A}\sin p(x_4+x_1),$$

which gives for the electromagnetic force

$$\mathrm{F}_{21} = \frac{\partial\kappa_2}{\partial x_1} = p\mathrm{A}\cos p(t+x) = \mathrm{F}^{21},$$

$$\mathrm{F}_{24} = \frac{\partial\kappa_2}{\partial x_4} = p\mathrm{A}\cos p(t+x) = -\mathrm{F}^{24}.$$

From which, $\mathrm{F}_{\alpha\beta}\mathrm{F}^{\alpha\beta} = 0$, and the energy-tensor $\mathrm{E}_\mu^\nu = -\mathrm{F}_{\mu\alpha}\mathrm{F}^{\nu\alpha}+\frac{1}{4}g_\mu^\nu\mathrm{F}_{\alpha\beta}\mathrm{F}^{\alpha\beta}$ becomes

$$\mathrm{E}_{11} = \mathrm{E}_{14} = \mathrm{E}_{44} = p^2\mathrm{A}^2\cos^2 p(t+x),$$

the other components being zero. Corresponding to equations (5), we now have $\mathrm{G}_{11} = -8\pi\mathrm{E}_{11}$, etc., instead of $\mathrm{G}_{11} = 0$. Picking out the three equations to be modified, it will be seen that (6*a*) is replaced by

$$h_{22}''+h_{33}'' = 16\pi p^2\mathrm{A}^2\cos^2 p(t+x),$$

and the other conditions are unaltered. Thus the periodic part of $h_{22}+h_{33}$ is

$$h_{22}+h_{33} = -2\pi\mathrm{A}^2\cos 2p(t+x). \qquad (20)$$

There is also an aperiodic part $4\pi p^2\mathrm{A}^2(t+x)^2$, indicating that the field is undergoing a cumulative change. This can only mean a secular change in the distribution of energy, etc., and, as is well known, waves of this kind do actually carry away energy from the source—not a small quantity of the second order negligible in our approximations as in the previous waves discussed, but a quantity of the first order.

Returning to the periodic waves (20), we notice that the gravitation potentials give no indication of the plane of polarisation of the waves. A varying electromagnetic potential κ_3 instead of κ_2 would produce just the same gravitational effect.

In order to determine h_{22} and h_{33} separately we have to find the non-electromagnetic wave $h_{22}-h_{33}$ emanating from the material source. The gravitational polarisation of light-waves thus depends on the disturbance of the uncharged material as well as on the electrical oscillations, and is quite unconnected with the electromagnetic polarisation.

When $h_{22}+h_{33}$ is not zero there exists a true Poynting vector $E_4{}^1$ representing a flow of true energy of the same order of magnitude as $h_{22}''+h_{33}''$. We shall examine whether there is not in addition a Poynting pseudo-vector $t_4{}^1$ of the second order of small quantities similar to that found in § 7. Equation (16) is not now applicable, but the general expression is (Eddington, 'Mathematical Theory of Relativity,' equations (59·4) and (58·52)).

$$16\pi t_\mu{}^\nu = -\frac{\partial}{\partial x_\mu}(g^{\alpha\beta}\sqrt{-g})\left[-\{\alpha, \beta, \nu\}+g_\alpha{}^\nu\{\beta\sigma, \sigma\}\right] \qquad (\mu \neq \nu).$$

When only h_{22} and h_{33} occur, this reduces to

$$\begin{aligned} 16\pi t_4{}^1 &= \frac{\partial}{\partial x_4}(g^{22}\sqrt{-g}) \cdot \{22, 1\}+\frac{\partial}{\partial x_4}(g^{33}\sqrt{-g}) \cdot \{33, 1\} \\ &\qquad -\frac{\partial}{\partial x_4}(\delta^{11}\sqrt{-g}) \cdot \{1\sigma, \sigma\} \\ &= -\tfrac{1}{4}\frac{\partial}{\partial x_4}(h_{22}-h_{33}) \cdot \frac{\partial}{\partial x_1}(h_{22}-h_{33})-\tfrac{1}{4}\frac{\partial}{\partial x_4}(h_{22}+h_{33}) \cdot \frac{\partial}{\partial x_1}(h_{22}+h_{33}). \end{aligned}$$

Accordingly the waves $(h_{22}+h_{33})$ carry away gravitational energy just like the waves $(h_{22}-h_{33})$. Light-waves by reason of their gravitational properties reduce the energy of their source by an amount slightly greater than is given by purely electromagnetic calculations. The ratio of this second order energy to the first order energy is usually very small indeed.

A New Method for Studying Ionising Potentials.

By H. D. SMYTH, Ph.D., National Research Fellow in Physics, U.S.A.

(Communicated by Prof. Sir Ernest Rutherford, F.R.S. Received August 9, 1922.)

Introduction.

It is evident that an exact knowledge of the energy necessary to remove the electrons from a molecule is of great value in formulating definite ideas of atomic structure and in testing the correctness of any quantitative theory. Not only is it desirable to know the work required to detach the outermost valence electron but also the further energy steps corresponding to the removal of more closely bound electrons. Moreover, in the case of polyatomic molecules, it would be most interesting to learn what the actual products of ionisation are, whether it is necessarily accompanied by dissociation and, if so, what is the state of the dissociated components.

In ionisation of a gas by the impact of slowly moving electrons we have a method of disrupting molecules under known energy conditions, and therefore turn to experiments of this type for the solution of the problems just propounded. The first important quantitative work was that of Franck and Hertz* in 1913 using a method originally due to Lenard. Modifications and improvements in this general mode of attack have enabled subsequent investigators to contribute greatly to knowledge of atomic structure. For a detailed discussion of methods and results reference can be made to Hughes' "Report on Photo-electricity"† or to a recent series of papers by Franck,‡ but for the purpose of the present paper it will be sufficient to state the principal results as briefly as possible.

It has been proved that there are various critical energies which an impacting electron can transfer to an atom, and that such a transference causes the atom either to eject one or more electrons entirely, becoming an ion or merely to change to a condition of higher potential energy, giving off radiation on its return to the normal state. In the cases of the monatomic gases and the metallic vapours the critical energies directly observed can often be interpreted in terms of electron transitions by comparison with spectroscopic experiments and the application of Bohr's theory of stationary states. Thus, the lowest ionising potential undoubtedly gives the energy necessary to remove one electron from the atom and agrees with the value from the known spectral series relations. For the higher degrees of ionisation, however, such indirect methods of interpretation become more difficult, and it is only in a few special cases that any success has been achieved. Similarly, in the case of the polyatomic molecules the complicating factor of dissociation is so troublesome and spectroscopic evidence so meagre that the theoretical interpretation of experimental results is usually doubtful.

---

* J. Franck and G. Hertz, 'Verh. d. D. Phys. Ges.,' vol. 15, p. 34 (1913).

† A. L. Hughes, 'Bull. National Research Council,' April, 1921.

‡ J. Franck, 'Phys. Zs.,' vol. 22, pp. 388–391, 409–414, 441–448, and 466–471 (1921).

The difficulty with the ordinary methods is that they are merely able to detect ionisation without giving any direct evidence as to what is being ionised or what the products are. Though it is sometimes possible by indirect means to distinguish between single and double ionisation or between molecular and atomic effects, it would be very advantageous to get a direct method of analysing the products of the ionisation tube.

Positive ray analysis as originated by Sir J. J. Thomson and developed by Aston and Dempster provides a method for determining the ratio of mass to charge, while the ordinary ionising potential experiments furnish a method for producing ions under known energy conditions. The combination of these two methods of analysis, therefore, should provide much more complete information. Unfortunately, such a combination presents serious difficulties. For the results of the positive ray analysis to be of any value, the mean free path of the ion must be comparable with the distance between the place of formation and its place of detection, since any collisions on the way might change the nature of the ion and thus invalidate the measurements of its formation energy. On the other hand, at the place of formation of ions the molecular and electronic densities must be sufficient to produce a detectable number of ionising collisions. Thus, the difficulty is to obtain the maximum number of molecules and electrons in one small region and as few as possible elsewhere. The ordinary arrangement of a hot filament and electrical fields confines the electrons within proper bounds and a unidirectional molecular stream such as is used in a mercury diffusion pump offers a possibility of localised high vapour density.

The experiments which will now be described are the first steps along the line of attack indicated in the preceding paragraphs. Mercury was selected as the most suitable substance for the initial experiments because it lends itself easily to the production of a unidirectional molecular stream and because its ionising potential is so well known. Its high molecular weight was also found of advantage in distinguishing its effect from that of light gaseous impurities.

Apparatus.

In discussing the apparatus, the main part of which is shown in fig. 1, the method of introducing mercury vapour will be described first. A continuation (not shown) of the tube above A, at the top of the drawing, was bent through 180° and led to the mercury reservoir which was heated electrically. A small tube B also led to the reservoir, but had a trap in it to prevent vapour from passing. Thus, mercury vapour coming from the reservoir entered at A into the region cooled by the circulation of water in the coil C. Much of it condensed on the walls, was caught in the trap T, and returned to the reservoir through B. A small portion, however, consisting of molecules whose thermal velocities were approximately parallel to the walls of the tube, passed down through T. Before getting to the second trap, U, many of these molecules diffused to the walls, but since the latter were cooled by freezing mixture in the cooling jacket, J, the probability of molecules rebounding from them was small. Consequently, the only molecules which were likely to pass through the trap U into the ionisation tube proper were those which passed

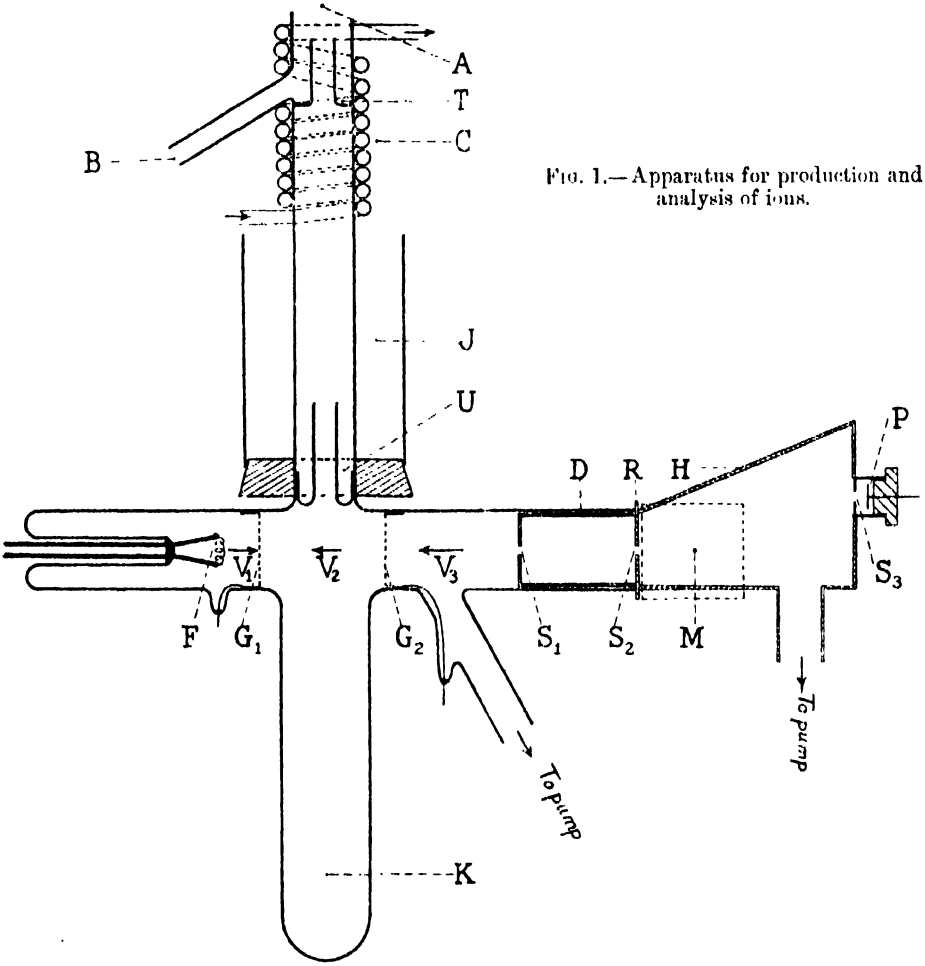


Fig. 1.—Apparatus for production and analysis of ions.

straight downward between the gauze electrodes, G_1 and G_2, into the condenser, K, which was immersed in liquid air. In this way general diffusion of mercury vapour throughout the apparatus was reduced to a minimum. Originally only the trap U was used, but it was impossible to arrange an outlet for it and it filled up rapidly. The introduction of a second trap with a return tube to the reservoir made it possible to use the apparatus continuously.

Through the side tubes shown, the apparatus was connected by ¾-inch tubing to a two-stage mercury condensation pump. The fore vacuum was produced by an oil pump in the early part of the work, but was later improved by the introduction of a Gaede rotary mercury pump. The pressure was measured on a McLeod gauge reading down to about 10^{-5} mm. Hg, and when the pumps were going there was no difficulty in keeping below this pressure, but there remained always a slight leak, presumably in the positive ray chamber, so that on standing overnight the pressure would go up to about 0·001 mm. A liquid air trap prevented mercury vapour from getting back to the apparatus from the gauge and pumps. Though the presence of wax joints made it impossible to bake out the whole apparatus, the mercury was first introduced in a supplementary reservoir so that the final reservoir could be baked out and then the mercury slowly distilled over into it under high vacuum.

The arrangements so far described furnished a fairly high and well localised density of mercury molecules between the platinum gauzes G_1 and G_2. We will now consider how some of these molecules were ionised, and the nature of the ions determined.

Electrons from the tungsten filament, F, were accelerated by the small field V_1 and then retarded by the equal field V_2, so that they would collide with the molecules between the gauzes at velocities between V_1 and zero (neglecting velocity of emission). The field V_2, besides preventing electrons from going beyond G_2, served to extract any positive ions formed. These were then further accelerated by the much larger field V_3 till they came to the slit S_1. The two slits, S_1 and S_2, of width 2 mm., served to define a beam of ions which then passed between the pole-pieces, M, of a small electromagnet, and were consequently deflected so that those of a particular value of m/e fell on the slit S_3 of width 3 mm. Behind this slit was the insulated plate, P, which was connected to a Compton electrometer kindly loaned to the author by the Department of Physics, at Princeton. The sensitivity of this instrument, as used, was about 4,000 divisions per volt.

The main part of the positive ray chamber, H, was of heavy brass, and was joined to the glass part of the apparatus at the flange, R, with sealing-wax. The cylinder, D, which extended into the glass tube, was of iron, so that the ions between S_1 and S_2 were shielded from the stray magnetic field. As the drawing of fig. 1 is not strictly to scale the most important dimensions are given below:—

Diameter of ionisation tube	2·2 cm.	
,, aperture in traps T and U	0·7 ,,	
,, tube containing traps	1·7 ,,	
Length of ,, ,, ,,	20 ,,	(Reduced in drawing.)
Diameter of condenser K	2·7 ,,	
Length of ,,	15 ,,	
Distance from F to G_1	0·4 ,,	(Exaggerated in drawing.)
,, ,, G_1 to G_2	3·5 ,,	
,, ,, G_2 to S_1	5·8 ,,	(Reduced in drawing.)
,, ,, S_1 to S_2	4·6 ,,	
,, ,, S_2 to S_3	8·5 ,,	
Area of poles of magnet	3·3 × 2·5 cm.	
Distance apart of poles of magnet	2·1 cm.	

The procedure adopted was to set the voltages, filament current, heating current, etc., and then measure the rate of charging of the electrometer as the current through the electromagnet was varied. Readings were always taken with increasing magnet current, and before each run the current was reversed several times. The magnetic field, corresponding to a given current, was read from a calibration curve obtained with a fluxmeter and coil. The electrometer readings were plotted against the magnetic fields so obtained.

After the apparatus was taken down the magnetic field was redetermined with a triangular coil, as described by Sir J. J. Thomson.* Within the limits of error these measurements gave the correct value of m/e for doubly-charged mercury ions, but a somewhat low value for the singly charged. This was presumably due to the fact that the coil used was not long enough to measure entirely the effect of the stray field, which was more marked at the higher values of the magnetic field required to bend the singly-charged ions. Incidentally, the stray field did valuable service by preventing any photo-electrons liberated at the end of the cylinder, D, around S_1, from getting back toward G_2 and causing spurious ionisation effects.

Results.

The wide slits used to get maximum intensity, together with the comparatively small magnetic field, made it impossible to get great resolution, but it was hoped that it would at least be possible to distinguish between singly- and

* J. J. Thomson, 'Rays of Positive Electricity,' p. 18, 2nd edition (1921).

doubly-charged mercury ions. It was found, however, that the effect of the single ions was so great, compared to that of the double at low voltages, that often the curves were of the type of A, fig. 2, showing both types of ion, but not completely separated. In curve B, on the other hand, where the value of V_1 was 80 instead of 40 and the cooling was better, the two effects are separated. In this curve there also appear two, and possibly three, other maxima of values of m/e approximately 15, 35, and 50, on the scale of $m/e = 1$ for hydrogen. These will be discussed in connection with a later curve.

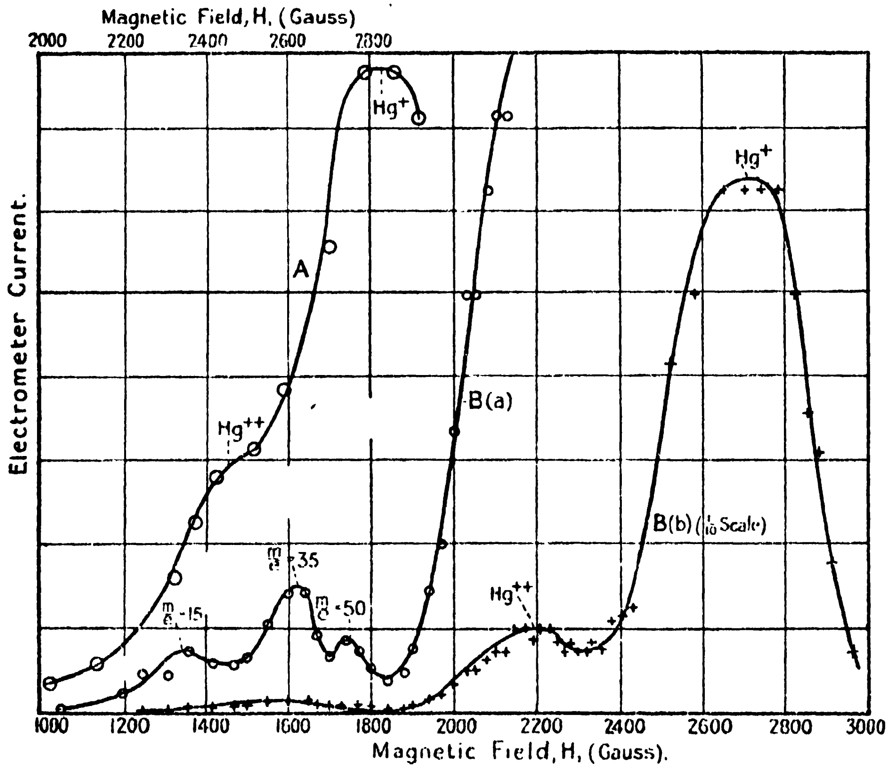


FIG. 2.

Curve A. $V_1 = -V_2 = 40$ volts. $V_3 = 620$ volts.
„ B. (*a*) $V_1 = -V_2 = 80$ volts. $V_3 = 520$ volts.
(*b*) Same run on 1/10 scale.

In fig. 3 a group of curves taken at higher values of V_1 is presented. In order to make readings possible these runs had to be made with very much reduced filament currents. The striking feature of these curves is the complete separation between the Hg^+ and Hg^{++} peaks, and the greatly increased relative intensity of the latter. In curves C and D again appear two peaks at low values of the magnetic field, H, and in D, where the maxima are unusually sharp, there is a small but definite maximum at $m/e = 65$ approximately. The effects at $m/e = 15$ and 30 are evidently due to the presence of a trace of air, and possibly other gaseous impurities. There was known to be a very small leak, and it was noticed that these peaks were always particularly strong when the pumps had just been turned on in the morning. Moreover, the addition of the Gaede pump in the fore vacuum greatly reduced their intensity, and it was found that on reducing V_1 they disappeared between 15 and 20 volts. The peak at $m/e = 65$ is naturally attributed to triply-charged mercury, and it is possible that mercury ions of higher charges are present but masked by the air peaks, just as it is difficult to identify with certainty

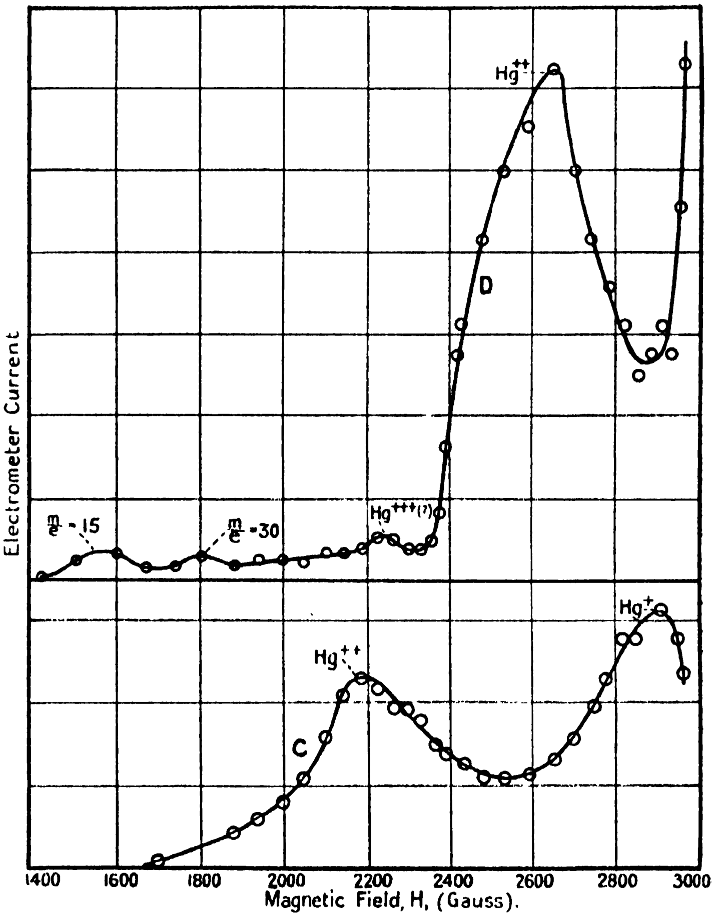


FIG. 3.

Curve C. $V_1 = -V_2 = 300$ volts. $V_3 = 300$ volts.
„ D. $V_1 = -V_2 = 175$ „ $V_3 = 610$ „

an $m/e = 65$ peak in the other curves. In any case, the number of more highly-charged ions is extremely small compared to that of the single and double, even when the voltage is run up to 500, and may be due to impacts among the ions themselves rather than to electron bombardment. It is, therefore, impossible to draw any conclusions as to the conditions for their formation from the present experiments, though it is hoped that a more powerful apparatus may be able to throw light on the question. Similarly, the search for a polyatomic mercury molecule must wait for a larger magnet.

The curves shown and discussed, so far, are typical of some sixty or seventy runs made with values of V_1 varying from 10 to 500, and V_3 from 120 to 1290. These served to give valuable evidence as to the production of singly- and doubly-charged ions, and to indicate roughly the beginning of double ionisation at between 20 and 25 volts. It remained to locate this point as definitely as possible.

It is clear that by the time the electrons get to the point of maximum mercury density between G_1 and G_2 they will have been slowed down by the field V_2, so that their speed will be considerably less than V_1, but cannot be properly estimated. The procedure adopted, therefore, was to make a run in the usual way, then set the magnet current at the value for the Hg^+ peak, and reduce the value of V_1 till the effect had completely disappeared. Since the ionisation potential of mercury is known to be 10·4, the difference between this and the observed voltage, where all effect disappeared, was taken as a correction to determine the effective voltage for the original run. Of course, this automatically includes corrections for initial velocity, potential drop across the filament, space charge, etc. When this procedure was first adopted curves were obtained of the type of K, fig. 4, giving corrections of 4 or 5 volts, which was the expected order of magnitude.

In the course of a week's running the magnitude of this correction gradually increased until it was as large as 15 volts in two or three of the last experiments. This may be connected with the gradual filling up of the trap U, and the consequent improvement in the cooling of the inner walls of the trap. In this way the number of stray molecules was reduced, so that the impacting electrons had to pass through a greater proportion of the field V_2 before colliding with mercury molecules. Notwithstanding this variation, the corrected results were very consistent.

In fig. 4 are presented the results of four runs taken at accelerating voltages of 13, 16, 20 and 25 (corrected), with corrections of from 3 to 5 volts. In curves E and F there is no trace of the Hg^{++} peak, while in G it is small but distinct, and in J unmistakable, though the Hg^+ peak is now spreading out. Of twenty-four curves taken at effective voltages below 25, the eleven taken at 20 or above, with a single exception, definitely show the Hg^{++} peak. The exception was taken at 20, and also shows a trace of Hg^{++} but not very definitely. A group of five runs at about 17 volts are conflicting, as two show Hg^{++}, two do not, and one is doubtful, but these were taken without liquid air on the condenser or ice in the cooling jacket, and therefore are not of great value. On the other hand, with proper cooling two runs at 16 volts, two at 15, and one at 13, show no Hg^{++}, while in one

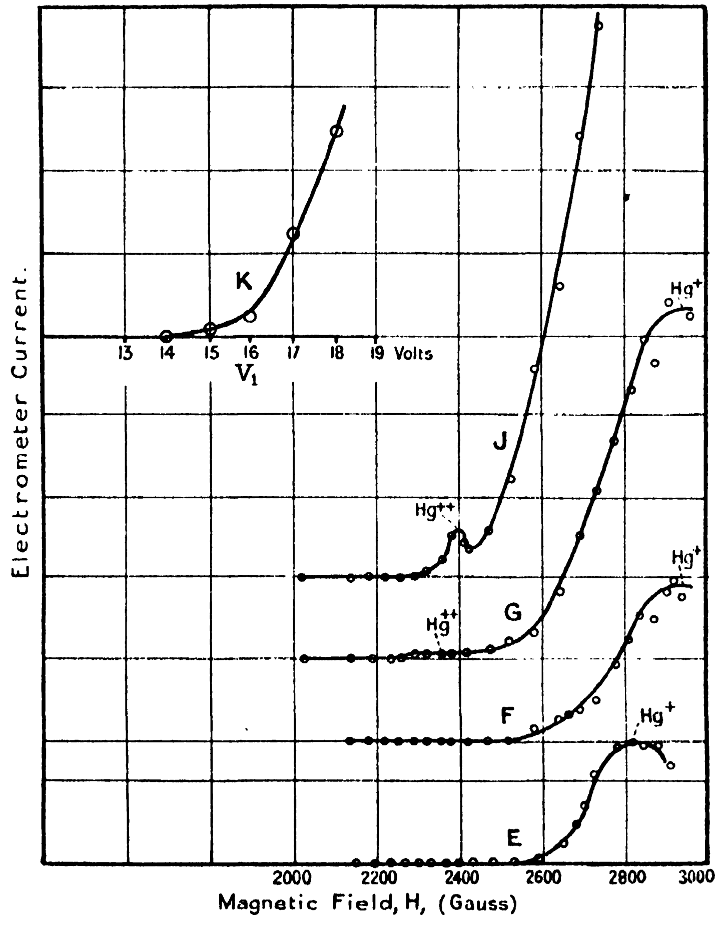


Fig. 4.

Curve E. $V_1 = V_2 = 13$ (17) volts. $V_3 = 450$ volts.
" F. $V_1 = -V_2 = 16$ (20) " $V_3 = 490$ "
" G. $V_1 = -V_2 = 20$ (25) " $V_3 = 490$ "
" J. $V_1 = -V_2 = 25$ (30) " $V_3 = 480$ "
" K. $H = 2820$ $V_3 = 460$. Correction curve for E.

Values of V_1 given are corrected. Uncorrected values in parentheses.

at 19·5 it is weak but definite, another at 19·5 does not show it, and a third curve at nineteen reveals just a trace of it. Three attempts to follow an Hg^{++} maximum to its vanishing point, with decreasing voltage, gave the results 22, 20 and 17 volts respectively. A consideration of these assembled results leads to the conclusion that doubly charged mercury ions are first produced when the electrons in the bombarding stream have attained an energy of 19 ± 2 volts.

Discussion of Results.

To appreciate the theoretical significance of the above results, it is necessary to review briefly some spectroscopic data. The enhanced spectrum of an element is supposed to be due to electron transitions in a singly ionised atom. Consequently, just as the principal series limit of the arc spectrum gives a basis for the calculation of the energy required to remove an electron from a neutral atom, so the corresponding series limit in the enhanced spectrum should give the energy necessary to remove an electron from a singly charged atom. This limit is not known for mercury, but has been determined approximately for zinc and cadmium, which are very similar to mercury spectroscopically. The limit of the principal doublet series of the enhanced spectrum of zinc is $\nu = 1\sigma = 147{,}544$, and of cadmium is $\nu = 1\sigma = 140{,}226$, where ν is the wave number.* Applying the quantum relation, $eV = h\nu c$, we obtain 18·2 volts and 17·3 volts for zinc and cadmium respectively. Now the single ion formation energies are 9·0, 9·4 and 10·4 volts for cadmium, zinc and mercury, and, by analogy, we should expect the double ion formation energy to be highest for mercury. Therefore the inference is that the energy required to remove a second electron from a singly charged mercury atom is of the order of 20 volts, and the energy necessary to eject two electrons at once from a neutral atom is $20 + 10{\cdot}4$, or approximately 30 volts.

If, now, we consider the experimental value of 19 ± 2 volts given above, we are forced to the conclusion that the double ions were formed first as the result of two impacts, but that there remains the possibility that higher speed electrons produced double ions at one impact. In the experimental curves the separation of the two maxima was not good enough to allow the detection of a sharp increase in the number of double ions in the neighbourhood of 30 volts. On the other hand, the curves at high voltages and low electron currents, *e.g.*, C, fig. 3, show almost as many double ions as single, indicating that, under such conditions, the majority of the double ions were produced by single impacts. This conclusion is supported by the fact that the ratio of the double ion effect to the single ion effect showed no consistent variation with the electron current.

---

* A. Fowler, 'Series in Line Spectra,' pp. 142, 146 (1922).

It is interesting to note that, in an attempt to detect double ionisation of mercury by a sharp increase in the total ionisation current, Einsporn* got kinks in his curves at about 17 volts and 42 volts. He attributes that at 17 to the presence of gaseous impurities, and is doubtful about the interpretation of the one at 42. In the light of the present experiments, it seems possible that the effect in the neighbourhood of 17 volts may have really been due to doubly charged mercury ions. The meaning of the 42-volt break remains doubtful.

It seems best to postpone any further discussion or speculation suggested by these results until further data are available. At present an attempt is being made to apply the method to gases, and a systematic investigation of the metallic vapours is reserved for the future.

Summary.

1. A new mode of attack on ionising potential problems is suggested. The principle involved is the study, by positive ray analysis, of the ions produced in a gas by the impact of slow-speed electrons of known energy.
2. The details of the application of the method to the study of mercury vapour are given.
3. Preliminary results indicate the formation of doubly charged mercury ions at 19 ± 2 volts. The series relations of the enhanced spectrum of mercury are not known, but analogy with zinc and cadmium suggests an estimate in agreement with the above value. The conclusion is that the double ions formed at this voltage are the result of two impacts. Experiments at higher voltages indicate formation by single impact.
4. More highly charged ions were present in such small quantities that their identification was uncertain, even at voltages as high as 500.

In conclusion, the author wishes to express his sincere thanks to Sir Ernest Rutherford for his continued and helpful interest and for the privileges of the laboratory, and also to Dr. Aston for some valuable suggestions. The work has been done while the author was a National Research Fellow in Physics.

---

* E. Einsporn, 'Zs. f. Phys.,' vol. 5, p. 218 (1921).

On the Analysis of α-Ray Photographs.

By P. M. S. Blackett, B.A., Charles Kingsley Bye-Fellow, Magdalene College, Cambridge.

(Communicated by Prof. Sir E. Rutherford, F.R.S.—Received October 5, 1922.)

[Plate 2.]

1. *Introduction.*

The beautiful method devised by Mr. C. T. R. Wilson* of making visible the tracks of ionising particles, by the condensation of water drops on the ions left in their trail, has been used in this investigation to study the passage of α-particles through gases. As is well known the tracks of these particles appear in these cloud photographs to be very nearly straight except for occasional sudden bends due evidently to collision with the nuclei of the atoms of gas. These bends occur very much more frequently towards the end of the range where the velocity of the particles is low.

The theory of the occurrence of such close collisions between the α-particles and the nuclei, based upon the assumption of a force varying inversely as the square of their distance apart, has been given by Rutherford† and Darwin‡ and verified by Geiger and Marsden,§ Chadwick‖ and others. These investigations confirmed in general the hypothesis but indicated a departure from the law of force when the two particles approach very close. Now it is also certain that at sufficiently great distances from the nucleus, the effective force on the particle must be different from that given by the inverse square law owing to the partial or total screening of the nucleus by the electrons. An investigation of the field at such greater distances can be done most sensitively by using particles of such a low velocity that they will not approach closer to the nucleus than the distance considered. As no particles have been detected with certainty with velocities less than about $0{\cdot}4\ v_0$, where v_0 is the initial velocity of the α-particles from radium *c*, by either the photographic or scintillation methods, it is necessary to make use of the Wilson expansion system.

The beginning of the study of the tracks of α-particles by this method has been recently described by Shimizu.* The present work, which has been carried out under the direction of Sir Ernest Rutherford, is a continuation of this and has already led to some interesting results. A statistical study of the frequency of the sudden bends in the tracks in both air and argon has shown that the velocity of the particles very close to the end of the track is as low as $v_0/20$, while the observed *distribution* of bends has been found to be consistent with the hypothesis that the inverse square law of force still holds between the nuclei and the particles when their distance apart is of the order of 10^9 cm. These results do not support the view that collisions of anomalous type and frequency occur towards the end of the tracks as has recently been suggested by Shimizu.†

---

* Wilson, 'Roy. Soc. Proc.,' vol. 87, p. 277 (1912).

† Rutherford, 'Phil. Mag.,' p. 669 (1911) and p. 538 (1919).

‡ Darwin, 'Phil. Mag.,' p. 500 (1914).

§ Geiger and Marsden, 'Phil. Mag.,' vol. 25, p. 604 (1913).

‖ Chadwick, 'Phil. Mag.,' p. 923 (1921).

In addition the study of the rare occasions on which an α-particle sets an atomic nucleus in rapid motion, giving rise to a branched track, has given interesting information about the dynamics of atomic collisions. The collisions considered appear to be perfectly elastic, no loss of energy or momentum being discovered. However this part of the investigation will be discussed in a later paper and our attention at present directed only to the statistical investigation of the deflections of the α-particles themselves.

2. *General Method.*

The procedure outlined by Shimizu has been followed closely in this work, and consisted in taking a large number of photographs of α-ray tracks, using a modified Wilson apparatus and a special camera, which by means of a system of mirrors gives on one film two images of the object viewed from directions mutually at right angles. The way in which the actual form of the track is calculated from measurements of the two images will be described. The result is a knowledge of the frequency of occurrence of bends of different angle at different distances from the end of the track, which is then compared with that to be expected theoretically.

Before considering this theoretical aspect, a short description will be given of the apparatus and its method of use.

3. *The Expansion Chamber.*

The expansion chamber used is one built for Mr. Shimizu, but with the reciprocating mechanism replaced by a simple spring action. This shares with Mr. Wilson's original design the advantage of obtaining very sudden expansions which can be repeated as frequently as required. The arrangement is shown in fig. 1. The speed of the expansion is regulated by the

---

* Shimizu, 'Roy. Soc. Proc.,' vol. 99, p. 425 (1921).

† Shimizu, 'Roy. Soc. Proc.,' vol. 99, p. 432 (1921).

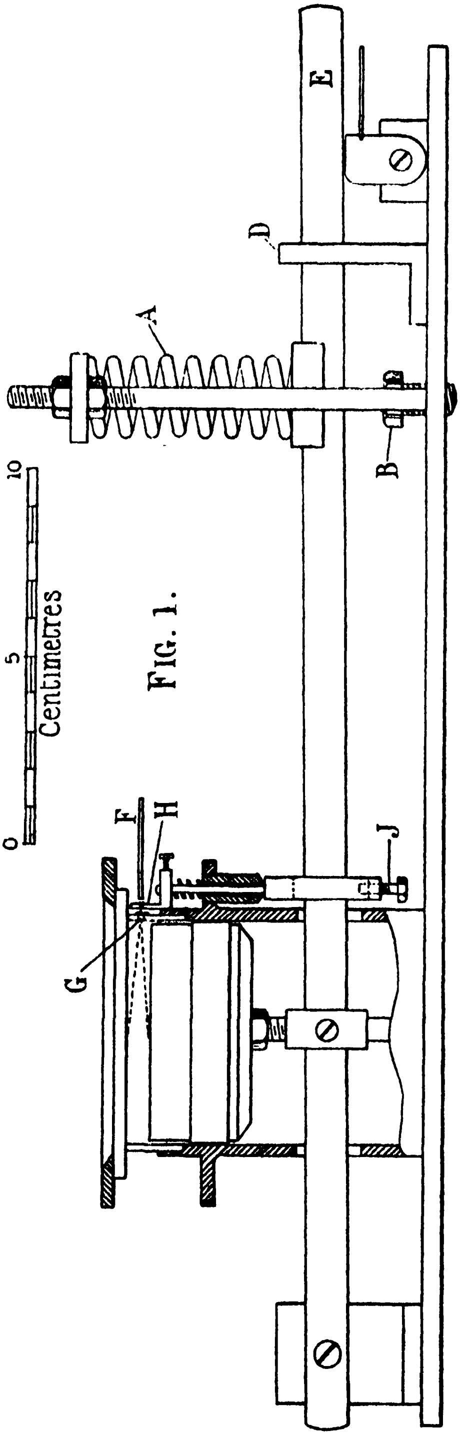


FIG. 1.

tension on the spring A and the amount of the expansion by the adjustable screw B. A paper scale on the bracket D, serves to measure the range of the bar E, and so the expansion ratio. The use of the expansion chamber and the coating of the internal glass surfaces with gelatine has been sufficiently described by Wilson and Shimizu. The source consists of a thin copper wire F, with a polonium deposit on its end. This is held in a carrier (not shown) opposite a small hole G in the glass cylinder wall, which is covered by a thin mica window of about 1 cm. stopping power. Between the source and the window is a mechanically operated shutter, H, worked by the main bar. An adjustable screw, J, allows the shutter to be so arranged that the stream of α-particles can only enter the chamber at the required moment, that is at or very near the bottom of the stroke. This keeps the chamber clear of all but sharp tracks. The stream of α-particles is restricted by means of suitable diagrams to a narrow cone, the axis of which is a diameter of the chamber. In order to facilitate the geometrical reduction of the photographs, the solid angle of this cone is kept as small as possible. The dotted lines in the figure show roughly the extent and position of the beam of rays.

The chamber is illuminated by two arc lamps placed symmetrically on either side, with suitable condensing lenses, heat-absorbing cells, and screens to limit the beams. By this means the two images in the camera are made equally bright.

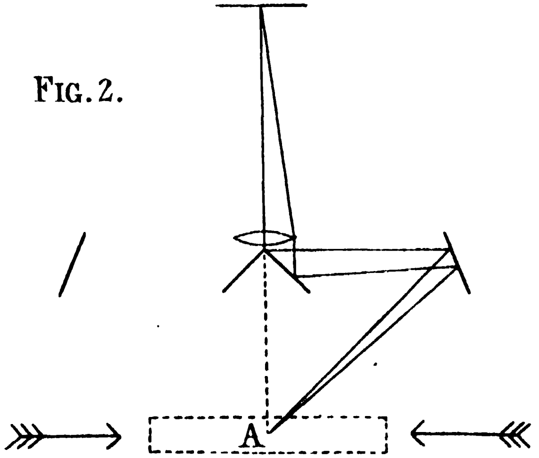
4. *The Camera.*

Shimizu (*loc. cit.*) described the optical arrangement of the camera, showing how the two images viewed from mutually perpendicular directions are projected through a single lens on to the film. His diagram is reproduced in fig. 2, but with the position of the arc lamps added so as to show the symmetrical arrangement. The camera is so placed over the expansion chamber that the line of intersection of the two focal planes is coincident with the axis of the cone of α-rays. That is, this axis would appear on fig. 2 as a line through A at right angles to the plane of the paper. The only modification of Shimizu's arrangement is the method of working the shutter. This is done by means of a Bowden wire, actuated by the main bar of the chamber, and adjusted so that the photograph is taken just as the expansion is completed. This has the advantage that the camera can be moved about as required and yet the shutter worked with the correct timing.

5. *Experimental Method.*

The apparatus is set up as has been described, and then the three adjustments made which are necessary to get good tracks. These are the expansion ratio, the speed of the expansion and the timing of the shutter which uncovers the source. This last should occur just before the bottom of the stroke. The speed of the expansion can be varied within fairly wide limits, but the expansion ratio must be very carefully adjusted. It was shown by Wilson*

FIG. 2.



that if v_1 is the initial and v_2 the final volume, then an expansion for which v_2/v_1 has the value 1·25 is necessary to produce the first rain-like drops in air, while a value 1·38 marks the beginning of the general fog. The value 1·33 was given as that which gives the best tracks. These results appear to apply to all diatomic gases expanded so rapidly that the conditions can be considered as adiabatic.

Now Wilson concluded that the type of cloud obtained depended only on the supersaturation at the end of the expansion, and therefore on the value of the ratio of the temperatures θ_1 and θ_2. Now this ratio for a gas, the ratio of whose specific heats is γ, is given for an adiabatic expansion by the relation

$$\frac{\theta_1}{\theta_2}=\left(\frac{v_2}{v_1}\right)^{\gamma-1}.$$

We can calculate, using this relation, the expansion necessary to obtain a given type of cloud for any value of γ, using the values of v_2/v_1 for air given above. In this way we find that, for monatomic gases, expansion ratios of 1·14 and 1·21 are necessary to give the first drops and the fog respectively.

Now, as there is no convenient way of measuring directly the expansion in use, we use these values given by Wilson for air, to calibrate the scale on the bracket D, which measures the relative travel of the bar. In this way the ratio v_2/v_1 in any experiment is determined with sufficient accuracy. It should be noticed that the marked difference between the requisite values of

---

* Wilson, *loc. cit.*

v_2/v_1 for monatomic and diatomic gases, affords a means of estimating the proportions of a mixture. This was seen very prettily in one experiment with argon, when a leak in the expansion chamber, letting in air, caused a gradual change in the expansion necessary to give good tracks.

When good tracks have been obtained the camera is placed in position and as many expansions made and photographs taken as desired. The photographs are then all examined, and whenever a track is seen to make a sudden bend, the two images on the photograph are measured up by means of a low-power microscope fitted with an eye-piece scale and cross-wire. A diagrammatic picture of the photograph of such a track is shown in fig. 3. The two images are shown slightly separated as they are actually in practice.

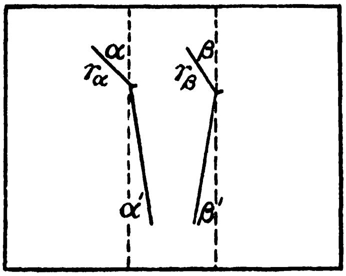


FIG. 3.

Since the information we desire from such a photograph is the angle ϕ of the bend and its distance r from the end of the track, the measurements we make are of the four angles marked α, β, α', β', and the two lengths r_α and r_β. The dotted reference lines from which the angles are measured are at right angles to the edge of the film, and are parallel to the image of the line of intersection of the two focal planes, which latter is the axis of the cone within which the stream of α-particles is projected.

It is now necessary to digress somewhat in order to show how from these six measured quantities we can obtain the actual deflection ϕ, and the distance r from the end of the track at which it occurred.

6. *Geometrical Theory.*

The two images are projections of the track upon two planes at right angles. If we call the line of intersection of these two planes Oz, and a line in each plane at right angles to Oz, Ox, and Oy respectively, we obtain a set of rectangular axes, from the origin of which we can draw two lines OP and OQ parallel to the actual directions of the initial and final parts of the track. Then the angles between OZ and the projections of OP and OQ on the Zx and Zy planes respectively, are the four angles α'; β', α, β which are measured on the photographs, as can be seen from fig. 4.

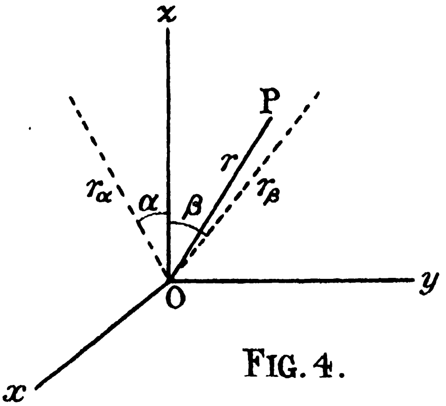


FIG. 4.

If the direction cosines of OP and OQ are l', m', n', and l, m, n, and if the angle POQ is denoted by ϕ, we find, using the relation

$$\cos\phi = ll' + mm' + nn',$$

and the four relations

$$l/m = \tan\alpha, \ldots \text{etc.},$$

that

$$\cos^2\phi = \frac{(\tan\alpha\tan\alpha' + \tan\beta\tan\beta' + 1)^2}{(\tan^2\alpha + \tan^2\beta + 1)(\tan^2\alpha' + \tan^2\beta' + 1)}, \tag{1}$$

which is the expression we require to calculate ϕ.

To calculate the length r of the line OQ, of which the projections on the planes Zx and Zy are r_α and r_β, we have

$$r^2 = r_\alpha{}^2 + r_\beta{}^2\sin^2\beta = r_\beta{}^2 + r_\alpha{}^2\sin^2\alpha, \tag{2}$$

which give

$$r = k_\alpha r_\alpha = k_\beta r_\beta, \tag{3}$$

where

$$k_\alpha = \frac{(1 - \sin^2\alpha\sin^2\beta)^{1/2}}{\cos\beta}; \quad k_\beta = \frac{(1 - \sin^2\alpha\sin^2\beta)^{1/2}}{\cos\alpha}. \tag{4}$$

The equations (3) and (4) are very convenient, since the values of k_α (or k_β) can be read directly from a set of curves of constant k_α drawn for all values of α and β. And this is the method always used, except for the cases when

$$\alpha = \beta = \pi/2,$$

when equation (4) becomes indeterminate and resort has to be made to equation (2).

Returning to the calculation of ϕ, it will be seen that equation (1) is of an inconvenient form, especially when a large number of tracks have to be examined, and although it is used in this investigation when special accuracy is desired, two other methods have been devised to simplify the calculations.

The first of these is based on the fact that, owing to the experimental arrangements, the angles α' and β' are small, except on the comparatively rare occasions when the bend being considered is subsequent to another bend of large angle.

Now for

$$\alpha' = \beta' = 0$$

(1) reduces to

$$\tan^2\phi = \tan^2\alpha + \tan^2\beta, \qquad (5)$$

while, when all four angles are *small*, ϕ is given approximately by the relation

$$\phi^2 = (\alpha - \alpha')^2 + (\beta - \beta')^2,$$

which to the same order of accuracy can be written

$$\tan^2\phi = \tan^2(\alpha - \alpha') + \tan^2(\beta - \beta'), \qquad (6)$$

which is of the same form as (5), with $(\alpha - \alpha')$ and $(\beta - \beta')$ written for α and β.

It is possible to investigate analytically the errors involved in using (6) instead of (1), but it is simpler to compare the two empirically in a number of cases. This has been done, and it has been found that this error is small when α' and β' are small, provided that α and β are not larger than about 45°.

Equation (6) is convenient, since ϕ is given as a function of the two variables $(\alpha - \alpha')$ and $(\beta - \beta')$, and can therefore be obtained graphically. For this reason this method has been employed in all cases when the conditions given above are fulfilled.

Now there is yet another way in which ϕ can be obtained, which is of special importance when α and β are equal to 90°; for in this case equation (1) becomes indeterminate, and the values of r_α and r_β have to be utilised to calculate ϕ. A sphere is ruled, with two intersecting sets of meridians at regular intervals, of perhaps 10°, from poles which are 90° apart. A photograph of such a sphere is shown on Plate 2, and the diagrammatic fig. 5 shows at once the method of use. In this figure the two polar diameters are called the axes of x and y, and the diameter at right angles to these the axis of z. Then any direction OP can be specified by a point P on the surface of the sphere, and this point P lies at the point of intersection of the meridians marked α' and β'. Similarly, any other direction OQ is specified by the two meridians α and β. And the angle POQ is given by the length of the arc PQ, which can be directly measured.

This method is very simple to use, and has the considerable advantage that it reveals at once the order of accuracy of the angle ϕ when obtained from any given set of measurements. For this will greatly depend upon the particular values of the four angles measured. An approach to the condition when ϕ is indeterminate will be revealed by the decrease of the angle between the two meridians which fix the point. For the less this angle, the greater clearly the indefiniteness in the position of the point.

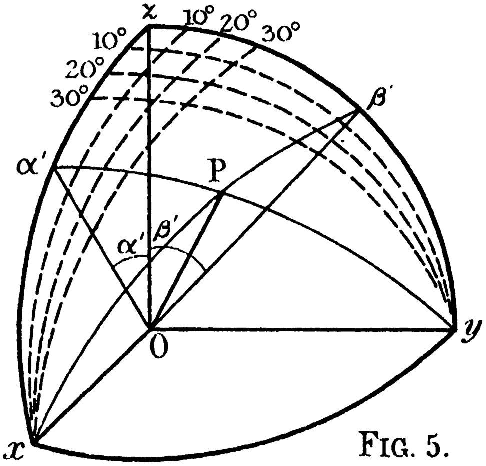


FIG. 5.

In this case of complete or approaching indeterminateness of ϕ, as calculated from the four angles alone, the observed value of r_α/r_β can be used in addition to fix the point Q. In fact, the sphere can be graduated in addition with curves of constant r_α/r_β, since this quantity is a function of α and β only. If this is done,* *all six* observed quantities can be utilised in the determination of ϕ, and suitable *weight* can be given to each of the variables.

The sphere was actually used in the *statistical* work on all occasions when the conditions for the use of equation (6) were unfulfilled.

For the suggestion of thus using the surface of a sphere to simplify the geometrical calculations, I am indebted to Mr. D. R. Hartree, of St. John's College.

7. *The Method of Using the Results of the Experiments.*

We obtain, as a result of these measurements and the subsequent calculations, a knowledge of all the noticeable bends made by a given number of α-particles. Enlarged reproductions of some of the photographs obtained are shown on Plate 2. Numbers 2 to 5 are of tracks in argon, and numbers 6 to 11 in air. The identification of the two images of a given track is assisted by noting that the z co-ordinate of any point, that is the distance from the bottom of the photograph, is the same in both images.

In particular, the end of both images of any track will be found on the same horizontal line. A good example of this is seen in number 6, in which five tracks appear, all ending at different distances (marked by arrows) from the bottom of the photographs. In the description of the photographs, bends will be described by writing the appropriate values of ϕ and r in brackets, ϕ being measured in degrees and r in divisions. Thus (20°, 42) means a bend of 20° at a distance of 42 divisions from the end of the track.

---

* The sphere shown in the plate is not provided with these graduations.

The simplest way of representing these results is to mark a point for each bend on a diagram on which ϕ and r are measured along two axes at right angles. In fig. 6 is shown a reproduction of part of such a diagram representing the results of the experiments with argon.

Certain cases will arise when two or more bends are made by a single α-particle, and in this case, strictly speaking, the deflection of the particle at any but the last of these bends should be plotted against the range that the particle would have travelled if the subsequent collision had not been made. But it will be seen from the theoretical discussion that it is legitimate to neglect such a complication. But it must be remembered that not the whole of such a diagram is physically significant, since bends of small angle, though relatively frequent, are not noticed. Thus the density of dots falls off rapidly for small value of ϕ, and as small angle bends are most easily noticed and measured at some distance from the end of the tracks, the limit ϕ_m, above which the observed distribution can be said to have any physical significance, will be smaller for the higher values of r.

Towards the end of the track, where bends occur fairly frequently, an appearance of general curvature is sometimes found. In most cases such an apparent curvature can be resolved into a succession of sharp bends, but even when this is not the case, it is still usually possible to decide whether or no a track has made a bend greater than some given angle.

It is now necessary to derive expressions for the distribution to be expected theoretically.

8. *The Theory of the Collisions.*

Rutherford and Darwin have given the theory of the frequency of collisions between α-particles and the nuclei of atoms, assuming the inverse square law and the conservation of kinetic energy and momentum during a collision. In order to adapt this theory to the present case, it is also necessary to assume that an α-particle of given velocity will travel a definite distance before losing its ionising power. This means that for all α-particles the remaining range r and the velocity v are connected by the same relation,

$$v = g(r).$$

Concerning this last assumption, it has been shown by both Darwin* and

---

* Darwin, 'Phil. Mag.,' p. 901 (1912).

Flamm* that the loss of energy of an α-particle to the nuclei of the atoms through which it is passing is on the average small compared to the loss to the electrons. The probability of variations in the loss of energy to the electrons has been calculated by Bohr† and Flamm, and shown to be small. Thus, on the average, the α-particles should behave sufficiently similarly to justify the assumption.

Though occasionally large angle bends do occur, in which a large proportion of the energy is lost to a nucleus, these cases do not appreciably affect the average behaviour of the particles. This assumption, however, is made less sure by the results of some recent work on straggling, which will, however, be discussed later.

Let M, E, v and m, e, u be the mass, charge and final velocity of the α-particle and nucleus respectively. Then if V be the initial velocity of the α-particle and ϕ and θ the deflection of the α-particle and the angle between the direction or the final motion of the nucleus and the initial direction of the α-particle respectively, then we have

$$\begin{aligned} MV &= Mv\cos\phi + mu\cos\theta, \\ 0 &= Mv\sin\phi - mu\sin\theta, \\ MV^2 &= Mv^2 + mu^2. \end{aligned}$$

From these we derive, following the analysis of Darwin, but writing η for M/m,

$$v/V = f(\phi), \tag{7}$$

and

$$\tan^2\theta = \frac{m^2}{(M+m)^2}F(\phi), \tag{8}$$

where

$$f(\phi) = \frac{\eta\cos\phi \pm (1-\eta^2\sin^2\phi)^{1/2}}{1+\eta},$$

and

$$F(\phi) = \operatorname{cosec}^2\phi\,[\cos\phi \pm (1-\eta^2\sin^2\phi)^{1/2}]^2.$$

If p be the distance from the initial position of the nucleus to the line of approach of the α-particle, Darwin showed that the calculation of the orbit gives

$$p = \tan\theta\,\frac{eE}{V^2}\left(\frac{1}{M}+\frac{1}{m}\right). \tag{9}$$

If there are N atoms of gas in unit volume, the probability, n, that a single particle will make a collision for which ϕ is less than some given value, when travelling a distance dr, is given by

$$n = \pi p^2 N dr. \tag{10}$$

---

* Flamm, 'Wien. Ber.,' IIa, vol. 123, p. 1393 (1914).

† Bohr, 'Phil. Mag.,' vol. 30, p. 581 (1915).

In such a collision the particle is deflected through an angle greater than ϕ.

From (9) and (10) we get

$$n = \pi N dr \frac{e^2E^2}{V^4}\left(\frac{1}{M}+\frac{1}{m}\right)^2 \tan^2\theta,$$

which using (8) leads to

$$n = \frac{K}{V^4}F(\phi)dr, \tag{11}$$

where

$$K = \pi N e^2E^2/M^2.$$

At this point Darwin's analysis becomes inapplicable to the present case, which requires the following procedure.

Differentiating (11) we get

$$dn = \frac{K}{V^4}F'(\phi)\,d\phi\,dr, \tag{12}$$

which gives the probability that an α-particle of velocity V makes a collision for which ϕ lies between ϕ and $\phi+d\phi$ when travelling a distance dr. In any collision of this type the particle will have its velocity reduced in the ratio v/V, but since the total number of collisions in which this ratio is appreciably less than unity is very small, the average velocity of the α-particles over the distance dr can be considered as unaltered.

Thus from (7) and (12) we obtain

$$dn = \frac{K}{v^4}F'(\phi)[f(\phi)]^4 d\phi\,dr. \tag{13}$$

This gives the probability of a collision in which the *final* velocity is v, and so after which the particle has a fixed remaining range r, given by

$$v = g(r).$$

Thus (13) becomes

$$dn = K\,F'(\phi)[f(\phi)/g(r)]^4 d\phi\,dr,$$

and the average number $\bar{P}$ of α-particles which are deflected through angles lying between ϕ_1 and ϕ_2 with remaining ranges lying between r_1 and r_2 will be given by

$$\bar{P} = QK\int_{\phi_1}^{\phi_2}\int_{r_1}^{r_2} F'(\phi)[f(\phi)/g(r)]^4 d\phi\,dr, \tag{14}$$

where Q is the total number of particles considered.

It will be seen that $\bar{P}$ is also the average number of dots which may be expected to be found in a rectangular area or cell of the distribution diagram limited by values ϕ_1, ϕ_2, r_1, r_2. From (14), supposing the right-hand side to be completely known, we can find $\bar{P}$ for all the n cells into which it is convenient to divide the distribution diagram. The actual numbers found in these cells can be counted. A consideration of the relation to be expected between the average distribution given by the theory and the actual observed distribution, which will of course be liable to chance variations, will occupy the next section of this paper.

9. *The Chance Variations in the Frequency of Bends of Given Type.*

It should be noted that the physical meaning of the average number $\bar{P}_r$ in the rth cell is that *average* number which would be found in that cell, in a very large number of separate experiments, all with the same total numbers of tracks, and all conducted under such circumstances that the theoretical laws of collision are obeyed.

Since we are supposed to be dealing with a large number of *dots*, we can consider the total number, S, to be the same in each experiment, in spite of the fact that the number is actually liable to chance variations.

Then all sets of numbers $P_1, P_2, P_3, \ldots P_n$, consistent with the equation

$$\Sigma P_r = S,$$

give possible, but not equally probable, distributions.

I am indebted to Mr. C. G. Darwin for having shown that the mean square departure of the observed numbers P_r for the rth cell from the average number $\bar{P}_r$ is, when all possible distributions are considered, given by

$$\overline{(P_r-\bar{P}_r)^2} = \bar{P}_r\left(1-\frac{\bar{P}_r}{S}\right).$$

If now we divide by $\bar{P}_r$, and sum over all the n cells, we find that

$$\Sigma\,\frac{\overline{(P_r-\bar{P}_r)^2}}{\bar{P}_r} = \Sigma\left(1-\frac{\bar{P}_r}{S}\right) = n-1,$$

since

$$\Sigma\,\bar{P}_r = S.$$

If then we write for every distribution,

$$F = \Sigma\,\frac{(P_r-\bar{P}_r)^2}{\bar{P}_r},$$

the *average* value of F for all possible distributions is given by

$$\bar{F} = n-1.$$

If, then, for any given distribution F has a value not far from $(n-1)$, we conclude that the distribution in question is an average one. But if F has a value greater than, say, $3n/2$, then the observed distribution is improbable and differs markedly from the average. But for physical purposes we can interpret a value of F about equal to $(n-1)$ as indicating that the *theoretical* *distribution utilised is the right one*, while a markedly larger value indicates that the *theoretical distribution is wrong.* The value of F thus appears as a criterion for the correctness of the theoretical average distribution.

However, it is not possible to calculate F for the observed distribution, since the right-hand side of (14) contains the unknown function $g(r)$. But the integration with respect to ϕ can readily be carried out, giving

$$\bar{P} = QKS(\phi_1, \phi_2)\int_{r_1}^{r_2} [g(r)]^{-4}\,dr, \tag{15}$$

where

$$S(\phi_1, \phi_2) = \int_{\phi_1}^{\phi_2} F'(\phi)\,[f(\phi)]^4\,d\phi.$$

We can now compare the observed and expected distributions by confining our attention to an area of the distribution diagram bounded by two values, r_1 and r_2. This area will cover a range of ϕ from 180° to ϕ_m, which latter will mark the end of the significant part of the diagram.

Equation (15) gives the relative numbers to be expected in each of the cells into which this area can be divided. Their absolute values can be found approximately by putting the expected number, $\bar{P}$, for the *whole area* equal to the observed number, P. If P is large this can be done without any large fractional error. Then the expected numbers for each cell can be calculated from (15), and the value of F for the distribution obtained. If this is not much greater than $(n-1)$, we conclude that the theoretical distribution is right. This being so, it is legitimate to proceed and calculate the velocity of the α-particle.

10. *The Calculation of the Velocity.*

Equation (15) gives for the area under consideration

$$\int_{r_1}^{r_2} v^{-4}\,dr = \frac{P}{QKS(\phi_m, \pi)},$$

where, as before, P has been written for $\bar{P}$.

If y be written for v^{-4}, and if $\bar{y}$ is its mean value over the range r_1 to r_2, then

$$\bar{y}(r_2-r_1) = \frac{P}{QKS(\phi_m, \pi)}, \tag{16}$$

and it is clear that there will be some value, r_0, between r_1 and r_2 for which y has its mean value. This value will be given by the solution of the equation

$$\int_{r_1}^{r_2} [g(r)]^{-4}\,dr = [g(r_0)]^{-4}[r_2-r_1]. \tag{17}$$

The procedure, then, for obtaining the function $g(r)$ from the distribution diagram is to evaluate $\bar{y}$ for each separate area, using (16), and by plotting the values so obtained against the values of $\frac{1}{2}(r_1+r_2)$ to obtain an approximate expression for $g(r)$. This is then used in (17) to give approximate values for r_0, the use of which instead of $\frac{1}{2}(r_1+r_2)$ gives a second approximation to $g(r)$. This process can be repeated as is found necessary.

It is interesting to note that in the special case of $g(r)$ being linear, that is of the form $(\mathrm{A}+\mathrm{B}r)$, the solution of (17) is

$$r_0'^4 = \frac{3r_1'^2 r_2'^2}{r_1'/r_2' + r_2'/r_1' + 1}, \tag{18}$$

where $r' = r + \mathrm{A}/\mathrm{B}$, etc.

This simple result, which is very useful as a first step in the approximation, shows at once that r_0 always lies on the low velocity side of the mid-point of the area. For from (18)

$$r_0' < (r_1' r_2')^{1/2} < \tfrac{1}{2}(r_1+r_2).$$

That this should be so is otherwise obvious.

In the actual calculation of the velocity for each area, it is convenient to measure r_1 and r_2 in divisions of the scale of the microscope, used in the measurements of the photographs. If there are q such divisions to 1 cm., and if p and M denote the expansion ratio and magnification of the camera respectively, while T and ρ are the initial temperature and pressure of the gas, it can be seen that the velocity is given by

$$v = \mathrm{K}_1\mathrm{K}_2 w,$$

where

$$\mathrm{K}_1{}^4 = \pi \mathrm{N}_0 e^2 \mathrm{E}^2/\mathrm{M}^2,$$

$$\mathrm{K}_2{}^4 = \rho \mathrm{T}_0/\rho_0 \mathrm{T} p q \mathrm{M},$$

and

$$w^4 = \mathrm{QS}(\phi_m, \pi)(r_2 - r_1)/\mathrm{P}.$$

And it will be seen that while K_1 depends on well-established atomic constants only, K_2 involves all the directly measured quantities, and w is given by the results of the statistical investigations.

The values of $\mathrm{S}(\phi_1, \phi_2)$, which are required for the experiments with air and argon now to be described, are given below. For the former, the value of η is 1/3·6 and the integration was done graphically. For argon, η is equal to 1/10, and the integration was carried out term by term, after expansion of the integrand in terms of ascending powers of η. It can easily be seen that for η equal to zero $(\mathrm{S}{\cdot}\phi_1, \phi_2)$ reduces to $(\cot^2\phi_1/2 - \cot^2\phi_2/2)$, which is the expression given by Rutherford. The values tabulated are for $\mathrm{S}(\phi, \pi)$, from which the values of $\mathrm{S}(\phi_1, \phi_2)$ can be obtained by subtraction. The values of $\cot^2\phi/2$ are given for comparison.

Table I.

	ϕ.	5°.	10°.	15°.	20°.	25°.	30°.	35°.	40°.	60°.
$\eta = 1/3 \cdot 6$	$S(\phi, \pi)$	524	121·5	47·4	27·0	16·4	10·4	6·8	4·4	1·1
$\eta = 1/10$	$S(\phi, \pi)$	524	127·1	54·8	29·7	18·1	12·0	8·5	6·1	2·1
$\eta = 0$	$\cot^2 \phi/2$	524	130·7	57·8	32·2	20·3	13·9	10·1	7·6	3·0

11. *The Experiments with Argon.*

The first actual experiments were carried out with a mixture of 75 per cent. argon and 15 per cent. nitrogen, since the high atomic number of argon, compared with air, results in many more bends being made by the α-particles. The distribution of bends made by the 281 tracks which were considered is shown on fig. 6. This diagram was divided for analysis into four areas, the

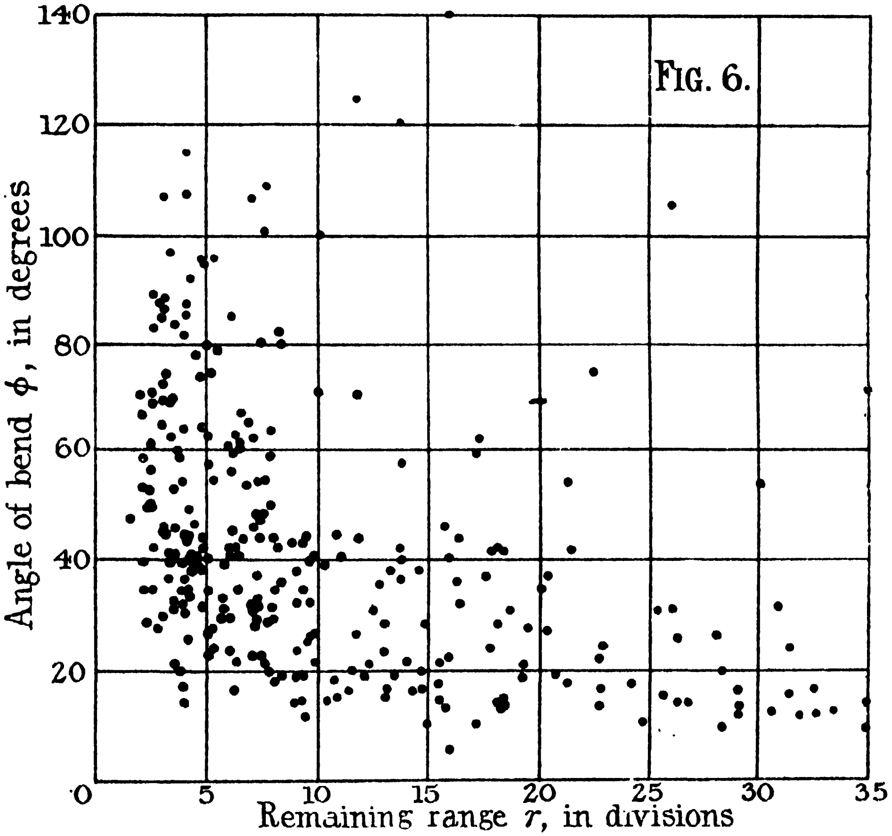


FIG. 6.

limits of which were 3 to 7, 7 to 15, 15 to 30, and 30 to 100 divisions.* Within each area the number of dots P between given limits ϕ_1 and ϕ_2 were counted, and are shown in Table II. Below each number P is the expected number $\bar{P}$, obtained as described above, by putting $\bar{P}$ for the whole area from π to ϕ_m equal to P for this area, ϕ_m being the limit of ϕ beyond which the distribution is clearly not in agreement with the theory. The values of P below this limit are enclosed in brackets.

Table II.

<table>
<tr><th colspan="2">ϕ_1, ϕ_2.</th><th>20, 30.</th><th>30, 40.</th><th>40, 50.</th><th>50, 60.</th><th>60, 70.</th><th>70, 80.</th><th>80, 90.</th><th>90, 180.</th></tr>
<tr><td rowspan="2">r_1 = 3
r_2 = 7</td><td>P ....</td><td>(11)</td><td>(20)</td><td>(22)</td><td>(8)</td><td>13</td><td>7</td><td>6</td><td>8</td></tr>
<tr><td>$\bar{P}$ ....</td><td>18·5</td><td>62·4</td><td>27·7</td><td>14·2</td><td>12·3</td><td>7·5</td><td>4·7</td><td>9·9</td></tr>
<tr><td rowspan="2">r_1 = 7
r_2 = 15</td><td>P ....</td><td>(21)</td><td>(17)</td><td>16</td><td>5</td><td colspan="3">7</td><td>5</td></tr>
<tr><td>$\bar{P}$ ....</td><td>71</td><td>23·6</td><td>10·4</td><td>5·4</td><td colspan="3">6</td><td>2·4</td></tr>
<tr><td rowspan="2">r_1 = 15
r_2 = 30</td><td>P ....</td><td>(12)</td><td>8</td><td>7</td><td>2</td><td colspan="4">5</td></tr>
<tr><td>$\bar{P}$ ....</td><td>32·8</td><td>10·9</td><td>4·8</td><td>2·5</td><td colspan="4">3·8</td></tr>
</table>

<table>
<tr><th colspan="2">ϕ_1, ϕ_2.</th><th>10, 20.</th><th>20, 30.</th><th>30, 180.</th></tr>
<tr><td rowspan="2">r_1 = 30
r_2 = 100</td><td>P ......</td><td>20</td><td>3</td><td>3</td></tr>
<tr><td>$\bar{P}$ ......</td><td>19·9</td><td>3·6</td><td>2·4</td></tr>
</table>

It will be seen that there is a rough agreement between the two sets of numbers. Further, since the value of F for the 16 cells to which the numbers relate is found to be 11·3, we can conclude that the observed distribution is consistent with that expected theoretically.

An objection may be made as to the arbitrary nature of the choice of ϕ_m. The justification for the procedure lies in the fact that the observed numbers in all cases agree with those expected for the higher values of ϕ, while they are invariably less for small values. This reduction in apparent frequency may be real, and due possibly to electronic screening, but many more experiments would be needed to establish it with certainty. The measurements on the actual photographs were carried down to a distance of about one-sixth of a millimetre from the end of the tracks, and their curlyness, so close to the end, makes it difficult to place much confidence in the number of *small* angle bends.

Since the observed and expected distributions are in agreement for values of ϕ greater than ϕ_m, we can proceed to calculate the velocities.

---

* 126 divisions correspond to 1 cm. equivalent range at N.T.P.

Putting $N_0 = 2{\cdot}705 \times 10^{19}$; $e = 17{\cdot}1\,\epsilon$;* $E = 2\epsilon$; $\epsilon = 4{\cdot}77 \times 10^{-10}$; $M = 6{\cdot}56 \times 10^{-24}$; we find that

$$K_1 = 1{\cdot}03 \times 10^8.$$

The values 760, 288, 1·22, 1/1·67 and 176, for P, T, p, M and q respectively give

$$K_2 = 0{\cdot}295,$$

while the values of w for the four areas are found to be 2·89, 4·81, 6·94 and 17·65 respectively.

If we now express the velocities as fractions of the initial velocity of α-particles from radium *c*, which will be taken as $1{\cdot}922 \times 10^9$ cm. per second, we obtain the following values for v/v_0.

Table III.

$\frac{1}{2}(r_1 + r_2)$	5	11	22·5	65
v/v_0	0·047	0·078	0·112	0·287
r_0	4·7	10·4	21·5	51·5
Range (in cm.)	0·037	0·088	0·17	0·40

In the first row is given the distance of the mid-point of each area from the end of the track, and in the second row the corresponding values of v/v_0. These are plotted as dots in fig. 7. The constants of the straight line which passes roughly through the points are used to evaluate r_0 for each area, according to equation (18). The last row gives these values the equivalent range reduced to centimetres in the gas at 15° and 760 mm. As circles in fig. 6 are plotted these final values of v/v_0 and r_0 and the smooth curve through the points represents the approximate form of the function $g(r)$ Owing to the limited number of tracks considered not much weight can be given to the actual curve, but the general magnitude of the velocity of the particles at these short distances from the end of the track is unmistakable.

---

* In the case of two mixed gases of atomic numbers z_1 and z_2 of which there are n_1 and n_2 atoms per unit volume, the total number of bends is proportional to

$$n_1z_1^2 + n_2z_2^2,$$

neglecting the correction due to the motion of the nucleus, which can be seen to be small. In these calculations the mixed gases were considered as simple gases of atomic number z_0 where

$$(n_1 + n_2)\,z_0^2 = n_1z_1^2 + n_2z_2^2.$$

For a mixture of 75 per cent. argon and 15 per cent. nitrogen $z_0 = 17{\cdot}1$.
For air $z_0 = 7{\cdot}25$.

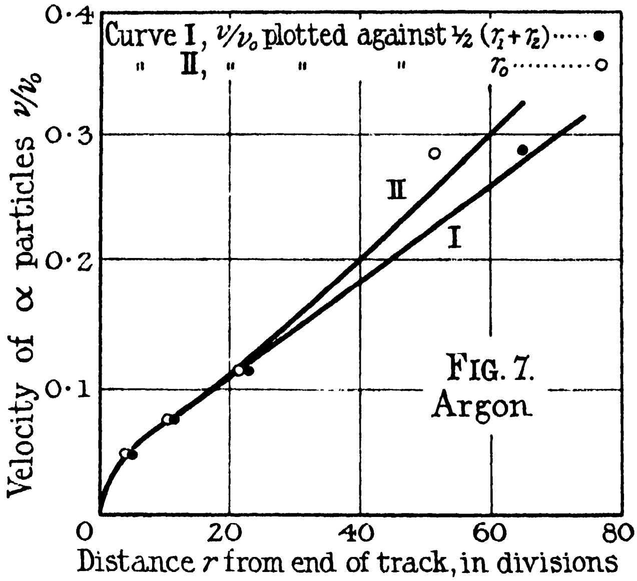


FIG. 7.
Argon.

12. *The Experiments with Air.*

The next work was carried out with tracks in air. The procedure was exactly as described for the argon tracks except that more tracks were observed, but not all of them measured up for all types of bend. The object was rather to determine the form of the function $g(r)$ than to verify the distribution with regard to angle over greater limits than were sufficient for the former purpose. Thus all bends down to the smallest observable angle were counted and measured for large values of r, while only bends over some larger angle were observed for the smaller values. The total number of tracks considered was 1524.

Putting $N = 2 \times 2{\cdot}705 \times 10^{19}$, and $e = 7{\cdot}25\,\epsilon$ we get $K_1 = 8{\cdot}08 \times 10^7$.

The measured values of P, T, p, M and q were 760, 288, 1·33, 1/1·62 and 153·3, which give $K_2 = 0{\cdot}295$.

The results of the statistical investigations are given in Table IV, and the final results of the experiment in Table V. In these Tables 126 divisions are equivalent to 1 cm. range at N.T.P.

Table IV.

ϕ_1, ϕ_2.	30, 40.	40, 60.	60, 180.
$r_1 = 3$ P	—	15	7
$r_2 = 5$ $\bar{P}$	—	16·5	5·5
$r_1 = 5$ P	12	12	5
$r_2 = 10$ $\bar{P}$	16·5	9·0	3·0

ϕ_1, ϕ_2.	15, 20.	20, 25.	25, 40.	40, 180.
$r_1 = 10$ P	15	10	9	5
$r_2 = 20$ $\bar{P}$	16·7	8·7	9·8	3·6

ϕ_1, ϕ_2.	10, 15.	15, 20.	20, 180.
$r_1 = 20$ P	11	4	7
$r_2 = 35$ $\bar{P}$	13·4	3·7	4·9
$r_1 = 35$ P	(9)	6	10
$r_2 = 60$ $\bar{P}$	25	6·9	9·1
$r_1 = 60$ P	13	11	5
$r_2 = 140$ $\bar{P}$	17·7	5·6	7·4

Table V.

r_1, r_2.	3, 5.	5, 10.	10, 20.	20, 35.	35, 60.	60, 140.
ϕ_m	40	30	20	10	15	10
P	22	29	30	22	16	29
Q	444	447	447	447	1524	1500
v/v_0	0·046	0·067	0·126	0·173	0·229	0·334
$\frac{1}{2}(r_1 + r_2)$	4	7·5	15	27·5	47·5	100
r_0	3·8	7·2	14·5	27·3	46·0	91
r_0	3·8	7·0	13·9	26·2	45·6	89
Range (in cm.)	0·0302	0·0555	0·111	0·208	0·362	0·707

The value of F for this distribution is 13·7 while $n-1$ is 16, so that we conclude that the theoretical distribution is the right one.

The points marked by dots on fig. 8 show the values of v/v_0 plotted against $\frac{1}{2}(r_1 + r_2)$. The curve through these points was taken as linear over the breadth of each area, and the value of A/B for each part used to calculate the values of r_0 given in the sixth row. The points given by the values of v/v_0 and r_0 were found to lie closely on a curve represented by

$$v = kr^{2/3}.$$

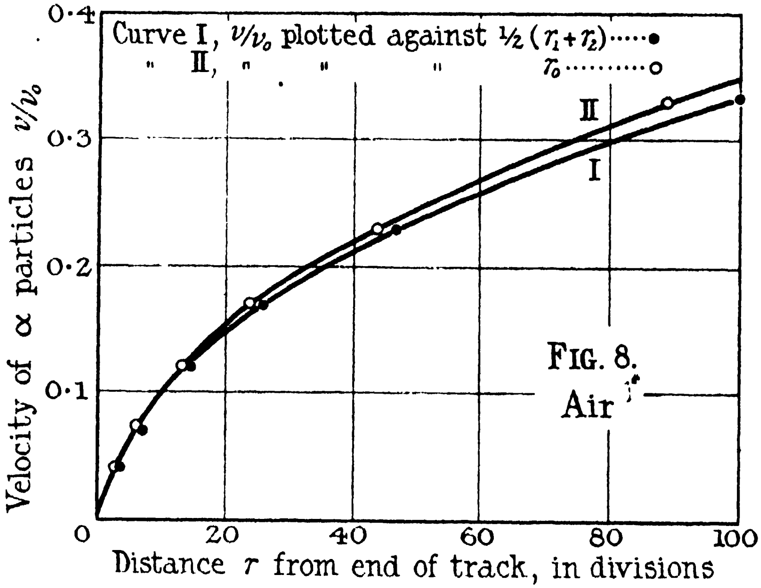


FIG. 8.

Using this expression in (17), the values of r_0 in the seventh row were obtained. The close agreement with the values in the previous row showed that the process of approximation had been carried far enough. The last row gives the final values of r_0 reduced to ranges in centimetres of gas at N.T.P. The points on fig. 7 marked with a circle represent the values of v/v_0 are plotted against the final values of r_0, and the comparison of the two curves indicates the difference between the first and the last approximation to the function $g(r)$.

13. *Discussion of Results.*

The agreement that has been found in these experiments between the observed and expected distributions allows us to deduce certain conclusions about the dynamics of the collisions.

In considering the question of the degree of verification given to the assumptions on which the theory is based, it must be remembered that there will be other average distributions appreciably different from that given here, which would be equally consistent with the observations. The latitude in the choice of such theoretical distributions will of course decrease as the number of observations increase, but with such a limited number as are here available, it is only possible to specify the *general type* of average distribution. Consequently, the only assumptions that can in any way be considered as verified are those to changes of which the theoretical distribution is sensitive.

Now it has been shown by Darwin* that such a distribution is sensitive to modifications of the law of force, the substitution of an inverse cube law, for instance, markedly altering the average distribution. Thus the observed agreement can be considered as indicating the existence of a law of force not much different from that of the inverse square. But it is less easy to decide to what degree the energy and momentum assumptions are verified. By far the most delicate test of these last is obtained by the study of individual branched tracks, into the consideration of which the law of force does not enter.

As to the question of the region in which the existence of inverse-square law is supported, it is necessary to calculate the maximum and minimum apsidal distances* of the orbits corresponding to the deflections considered. In the case of both argon and air, the most intimate orbits are those in which an α-particle of velocity about $0{\cdot}45\, v_0$ is deflected through an angle of about 30°, while the least intimate are those in which particles of velocity $0{\cdot}04\, v_0$ are deflected through 60° and 40° respectively. The apsidal distances in these cases are about $7{\cdot}4 \times 10^{-12}$ and $5{\cdot}4 \times 10^{-10}$ cm. for argon and $3{\cdot}2 \times 10^{-12}$ and $3{\cdot}7 \times 10^{-10}$ cm. for air. And since an orbit is appreciably determined by the forces at distances considerably greater than the apsidal distance, it can be said that these experiments support the hypothesis that the inverse-square law holds between the nucleus and an α-particle when their distance apart lies between about 7×10^{-12} and 10^{-9} cm. in the case of argon and 3×10^{-12} and 5×10^{-10} cm. in the case of air. It is interesting to note that the upper limit for argon is considerably greater than the distance of the K ring electrons from the nucleus. From the results, however, it appears possible that, especially in the case of argon, there is a departure from the inverse-square law at distances greater than those given.

It will be noticed that no confirmation of the magnitude of the nuclear change can be obtained from these results unless the velocities here calculated are independently confirmed. But the discovery that the velocities in air and argon are found to be almost identical close to the end of the range, in spite of the fact that bends in general are about three times more frequent in argon than in air, does indirectly confirm the assumption. For the range in argon of an α-particle from polonium has been shown by Adams† to be only 12 per cent. greater than in air, so that it would be expected that the relation between the velocity and the range would also be similar.

The analysis reveals the existence of α-particles of very low velocity compared with any that had been observed before. It appears that α-particles at a distance of about a quarter of a millimetre from the end of the track have a velocity about $0 \cdot 04 v_0$. Now it can easily be seen that this is just about the minimum velocity which must be possessed by an α-particle in order that it should impart to a free electron energy corresponding to a fall through 13·5 volts. In other words, the minimum velocity measured is roughly that at which the α-particle may be expected to lose its ionising power in air.

---

* Darwin, *loc. cit.*

† Adams, 'Phys. Rev.,' vol. 24, p. 113 (1907).

There is considerable difficulty in comparing these results with the data obtained by other methods, for the reason that these results refer to the average velocity of a number of α-particles at a given distance from the end of the track, while all other results refer to the average properties of a beam of α-rays at a given distance from the source. However, Marsden and Taylor* have shown that the function $g(r)$ is of the form $r^{1/3}$ when the absorbing material is air, and r is greater than about 2 cm., while both their work with heavy metals as absorbing material and the theoretical investigation of Bohr suggest that for lower velocities the speed of the particles will be proportional to some higher power of r.

The present results are fully in accord with this view since, as has already been mentioned, the relation

$$v = kr^{2/3}$$

appears roughly to represent the observations, which are all for values of r less than 1 cm. But it does not seem possible without further data to discuss usefully the transition between these two approximate impressions for $g(r)$. For it must be remembered that the statistical investigations are based on the assumption that an α-particle of given velocity will travel a given distance. If, as has recently been suggested by Henderson,† a large degree of straggling occurs in the last few millimetres of the range, this assumption will be only very approximately true. But though it is unlikely that the neglect of this phenomenon in the analysis can seriously affect the order of magnitude of the calculated velocities, it must certainly instil caution against attaching undue weight to the exact form discovered for the relation between the velocity and the remaining range.

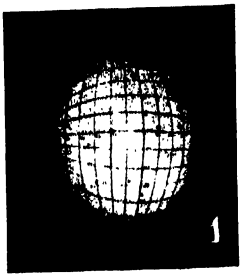
In conclusion, it may be said that the recent measurements by Kapitza‡ of the heating effect of a beam of α-rays give, for the average velocity at a given distance from the *end of the beam*, values nearly the same as those given by the statistical method for the average velocity at a given distance from the

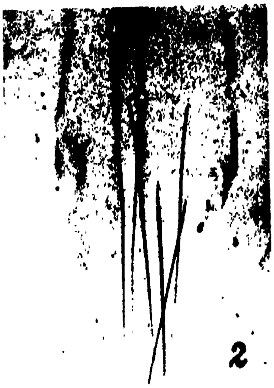
---

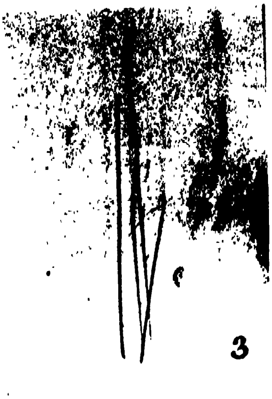
* Marsden and Taylor, 'Roy. Soc. Proc.,' A, vol. 88, p. 443 (1913).

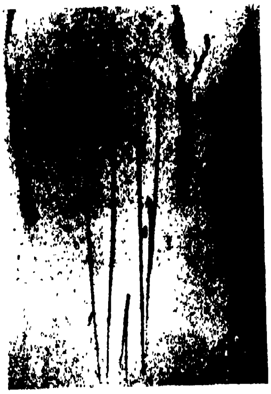
† Henderson, 'Phil. Mag.,' vol. 42, p. 538 (1921).

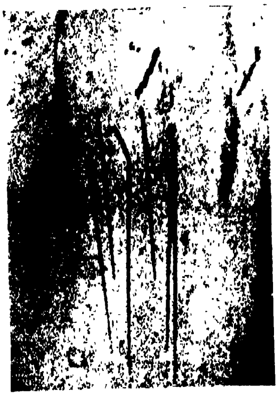
‡ Kapitza, 'Roy. Soc. Proc.,' vol. 102, p. 48 (1922).

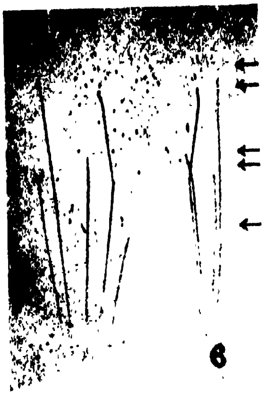

1


2

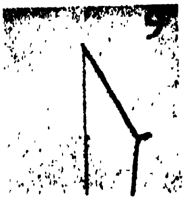

3

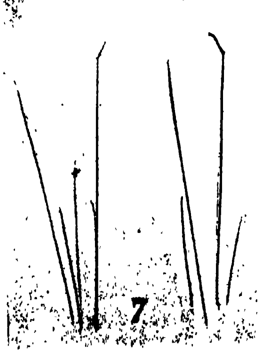


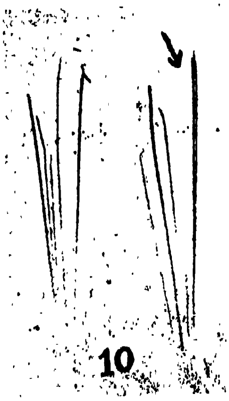


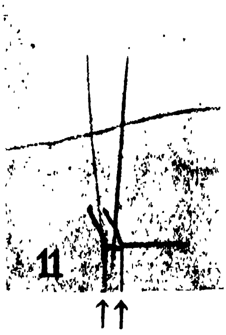

6

8


9


7


10


11

end of each track. The actual values are shown below, and refer to air at 15° and 760 mm. :—

Distance from source (in cm)	6·26	6·50	6·72	7·07
Distance from end of beam (in cm)	0·81	0·57	0·35	0
Heating effect (Kapitza)	0·400	0·316	0·244	0·084
Frequency of bends	0·365	0·300	0·224	0

If no straggling occurred this agreement would be expected In view of the large degree of straggling which does occur, it is a little difficult to account either for this agreement or for the disagreement which would be shown if the straggling were to be taken into account.

14. *Summary.*

1. Mr. C. T. R. Wilson's expansion method was used to study in detail the passage of α-rays through air and argon, attention being especially directed to the latter part of the range.
2. The frequency of large angle deflections was found from the study of the photographs of a large number of tracks, and compared with that to be expected on the nuclear theory of the atom.
3. The agreement found indicates that the inverse square law holds between the nuclei and the α-particle, when their distance apart lies between about 7×10^{-12} and 10^{-9} cm. for argon, and 3×10^{-12} and 5×10^{-10} cm. for air.
4. The velocity of the α-particle along the latter part of its range was determined from the frequency of bends and found to be as low as $v_0/20$ very near the end, where v_0 is the velocity of the α-particles from radium C.
5. No anomalous effect was discovered either as regards frequency or type of collision.

My thanks are due to Sir Ernest Rutherford for his continual interest and advice throughout this work.

DESCRIPTION OF PLATE 2.

All these reproductions are 3·06 times larger than the photographs

Nos. 2 to 5 are of tracks in argon, for which $p = 1·22$, $M = 1/1·67$. Hence these tracks are 2·24 times larger than they would be in the gas at 15°, 760 mms.

Nos. 6 to 11 are of tracks in air, for which $p = 1·33$, $M = 1/1·62$. Hence these tracks are 2·51 times as large as they would be at 15°, 760 mms.

In all the photographs 1 cm. corresponds to about 60 divisions of the microscopic scale.

No. 2.—The long track shows a bend (75°, 23) followed by a short hook near the end.

No. 3.—A single (109°, 8) bend.

No. 4.—A (63°, 17) followed by a (48°, 6) on the long track. The two images have their ends nearly super-imposed. Also a (140°, 16) lower down the photo. Note the short "spur" to this latter and compare it with the long spur due to the lighter "air" atom in No. 10.

No. 5.—A (31°, 26) followed by an (80°, 7).

No. 6.—One track showing an early bend (20°, 65) with a small spur.

No. 7.—A (42°, 19) bend also showing spur.

No. 8.—A (52°, 16) with a long spur. The left-hand image marked by an arrow shows that the three parts of the track are co-planar.

No. 9.—An early bend (39°, 18) with long spur. The unusual breadth of the r.-h. image is probably an optical effect.

No. 10.—A marked bend (132°, 18) indicating a nearly head-on collision. The r.-h. image is marked by an arrow.

No. 11.—A remarkable track showing an (895°, 59) bend. The direction of the initial part of the track upwards. The l.-h. image shows the remaining part of the track very much foreshortened. The long spur shows two marked bends, evidently due to collision of the recoil atom itself with two atoms of the gas. This track is the most suitable for accurate measurement, and for it the mass ratio m/M is found to be 3·3. Almost certainly then the atom collided with is a nitrogen atom.

On the Structure and Chemical Activity of Copper Films, and the Colour Changes accompanying their Oxidation.

By C. N. Hinshelwood, B.A., Fellow of Trinity College, Oxford.

(Communicated by Prof. J. W. Nicholson, F.R.S. Received August 30, 1922.)

When the surface of a metal is exposed to the action of a gas with which it reacts chemically, brilliant colour phenomena are frequently produced. In some cases the colours are recognised as diffraction colours, produced by the scattering of light in the surface film, and not as simple interference effects. When this is found the film must have a more or less complex structure, fine grained compared with the wave-length of light, but of a coarse-grained granular nature compared with molecular magnitudes. The work of Beilby has drawn attention from other points of view to the complex structure which the surface layers of metals, or thin films of metals, may assume.

It seems to be of considerable importance to correlate the chemical activity of such surface films with their structure, in view of the bearing this correlation may have on the problems of heterogeneous catalysis.

A first attempt in this direction is made in this paper, the reaction investigated being the oxidation and reduction of a copper—copper oxide film on the surface of metallic copper.

The general nature of the phenomena observed is as follows: When bright, rolled copper foil is exposed to the action of oxygen gas at low pressure, and at 200—300° C., a film of oxide forms on the surface, and, as oxidation proceeds, the velocity of oxygen absorption diminishes considerably, owing to the difficulty of penetration of the oxide film. The surface assumes various tints, all of which, however, are quite faint, finally becoming black. Its metallic lustre is retained throughout. After reduction, a second oxidation is brought about very much more readily than the first, and the surface assumes a new series of tints, brighter than the first, and even qualitatively distinct from them. Successive oxidations and reductions take place more and more readily until, ultimately, a limiting rate appears to be attained, which may be some hundreds of times greater than the original rate. As the surface becomes more and more active chemically the diffraction colours simultaneously increase in brilliancy. Ultimately a permanent colour sequence, with tints of great brilliancy, is established, which may be traversed time after time. The surface loses its metallic lustre, and when in the fully activated condition the copper is salmon-pink in colour.

During the activation process the structure of the surface film seems to change in the following way: At first the surface layers of the copper foil are compact as a result of its mechanical treatment; during the successive stages of activation by alternate oxidation and reduction, the copper atoms in the surface film are able, under the influence of surface tension, to aggregate themselves more and more completely into small discrete units in what Beilby calls "open formation," and the film assumes a granular structure freely permeable to oxygen. From the experiments to be described it is possible, in principle, to determine the size of these small granules, and although there are difficulties in the way of finding accurate values, it emerges clearly that the order of magnitude of their diameter is only a small fraction of 1μ. The view has been expressed that metallic grains as small as this assume a spherical form.* But Sir George Beilby has kindly pointed out to me that the evidence of microphotographs shows the film to be lenticular in structure. In the last section of this paper the magnitude of the granules is calculated, for the sake of simplicity, on the assumption that they are spherical. The radius thus determined will be their mean or "effective" radius.

Each small unit of copper will be oxidised independently, and the extent to which it is converted into oxide determines the colour of the diffracted light. The amounts of oxygen absorbed, corresponding to various well marked steps in the colour sequence, has been determined, and may, incidentally, be of interest in connection with the optics of the phenomenon.

---

* Maxwell Garnett, 'Phil. Trans.,' A, vol. 205, p. 279.

There is a curious difference between the phenomena observed during reduction and those observed during oxidation, which suggests an interesting mechanism for the gradual activation of the film.

Experimental Method.

The amounts of oxygen, or hydrogen, absorbed were measured by means of a McLeod gauge, so constructed that it could be used over the range of pressures corresponding with the passage of the surface films through their various stages of colour. Pieces of copper foil, of known area, were used, and very carefully cleansed.

The apparatus is shown in fig. 1. Bulb A contains the copper, and is kept at constant temperature by being surrounded with the vapour of a liquid of suitable boiling point. CD is the McLeod gauge in a convenient form, similar to that described by Merton,* the mercury being raised, or lowered, by letting air into, or evacuating D, as required. B is a trap. E is a volumenometric device, connected to the main apparatus at the end of a series of experiments.

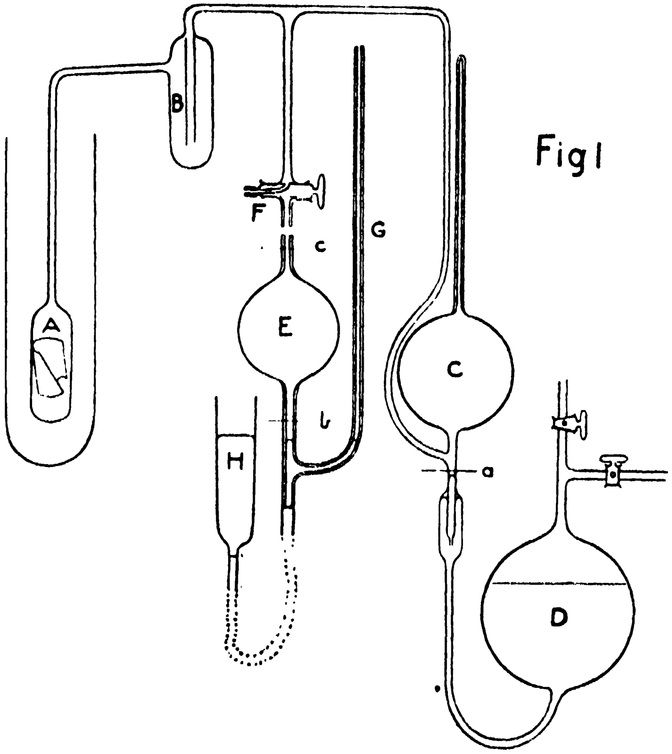


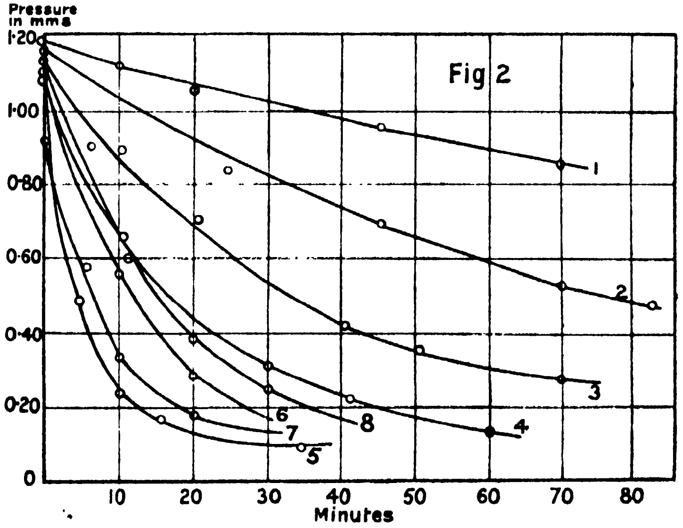
Fig I

In an experiment the apparatus is evacuated by means of the three-way tap F, filled once or twice with oxygen, or hydrogen, and then evacuated to the desired pressure, which was initially about 1 mm. in one series of experiments and 4 mm. in another. To start an experiment the vapour jacket was placed round A, and the diminution of pressure with time was observed. Except when a reading is being made the mercury is kept at the mark *a*. To convert pressure readings to absolute measure the total volume of the apparatus needs to be known. The bulb A, moreover, and part of the connecting tube are at a high temperature, and the rest at room temperature. To find the effective volume of the whole, corrected for this temperature difference, connection is made to E, and air admitted at atmospheric pressure. The increase of pressure necessary to compress from *b* to *c* is measured on the gauge G, the mercury in the McLeod gauge always being kept at *a* by applying pressure to D. The volume between *b* and *c* being known, the effective volume of the rest at the actual temperature of E is readily found.

---

* 'Roy. Soc. Proc.,' A, vol. 90, p. 550.

With this apparatus it is possible to make measurements with considerable precision, as may be seen from fig. 2, in which the experimental points are marked on the curves.



Successive oxidations at 233°.

In the reduction experiments the minute amounts of water vapour present do not influence the readings, since the actual pressure in the capillary of the gauge when a reading is made is greater than the saturation pressure.

Retention of hydrogen by the reduced copper might have been expected to cause trouble. Only once in the course of about fifty experiments was there any indication of such a disturbing factor, and even then there was no proof that it arose definitely from this cause. Probably any minute amounts of hydrogen retained are oxidised even before the initial reading is taken.

Apart from their intrinsic interest it is necessary to observe the colour phenomena carefully, for the following reason. The colours being diffraction and not interference colours, depend upon the structure, and not upon the thickness of the film. That is to say, they are independent of the depth to which the granular active film extends and are determined solely by the size and degree of oxidation of the units constituting this film. Below the film is compact metal. The active film oxidises very rapidly indeed compared with the compact metal below, so that in an experiment not unduly prolonged we may regard the chemical changes as confined to this film. In order to give an absolute significance to a series of measurements of oxygen absorption, we must know the quantity of metal which composes the film taking part in the reaction. If we know the weights of oxygen necessary for the oxidation of two films to the same colour, we obtain the relative amounts of metal in the two; and if we know the weight of oxygen necessary to oxidise a given film black, that is completely to cupric oxide, then the amount of copper in the film is known. Quantitative results on the very large diminution in velocity of oxidation as soon as the black stage is reached leave little doubt that this does actually represent complete oxidation of the activated film, and that subsequent action depends upon the slow attack of the massive metal beneath.

The Colour Phenomena.

That the colours shown by oxidised copper and by tempered steel are due to diffraction was pointed out by Mallock,[*] who found that a film of given tint could be ground away while retaining its colour, which must therefore be dependent not upon its thickness, as it would be if interference were the cause, but upon its minute structure. Corroborative evidence is cited by Raman.[†] The relation to Beilby's observations has also been pointed out.

All the quantitative results of these experiments bear this out fully.

1. In the oxidation of copper, the brilliancy of the diffraction colours increases *pro rata* with the chemical activity, the granular film which most effectively scatters light presenting the largest surface to the action of the oxygen. When fresh copper is oxidised the colours are faint and the sequence is not constant, but is usually silvery or steely, pale straw, violet, black.

The permanent colour sequence is purple, blue, green, very light green (almost yellow), purple, blue, black. These are very brilliant, except the second appearance of purple and blue; they appear uniformly and are quite constant in hue from time to time.

---

[*] 'Roy. Soc. Proc.,' A, vol. 94, p. 561.

[†] 'Nature,' January 26, 1922.

2. The absolute amount of oxygen absorbed corresponding to a given colour depends upon the amount of copper in the granular film, but the ratios of the amounts of oxygen required for the attainment of the various stages are constant. This shows clearly once more that the colours are structure colours. Thus for oxidation corresponding to maximum brightness, as far as the eye can judge, of the various colours, we have:—

	Grammes cupric oxide per 1 sq. cm. of foil.	Ratio taking light green as unity.	Grammes cupric oxide per 1 sq. cm. of same foil after heating for 10 hours in air and reduction.	Ratio.
Purple	0·000030	0·41	0·000067	0·38
Blue	0·000043	0·59	0·000115	0·64
Light green ...	0·000073	1·00	0·000178	1·00

The average of another series of experiments leads to the following set of ratios, taking the amount corresponding to light green as unity:—

Purple	0·40	Second purple..............	1·3
Blue	0·64	„ blue	1·6
Light green	1·00	Black	1·9

The oxidation of an active film must be imagined to consist in the independent oxidation of each of the small granules composing it, a layer of oxide being formed on the surface of each and extending inwards. The size and composition of each thus changes, and consequently the colour of the film alters. Thus, when the granules consist of approximately one-third oxide and two-thirds copper, the colour of the scattered light is bright blue. Within the boundary of each granular unit of copper or of copper oxide there is no reason to doubt that the density of the material is normal, the low mean density of the film being due to the spacing of the granules.

The effects observed during reduction differ in a remarkable manner from those observed during oxidation. No diffraction colours appear when a black film is reduced, the colour of the film at any stage of the reduction being simply a shade of brown, produced by the combination of the red colour of the copper and the black of the oxide.

At first it is natural to ask whether the difference can be due to the intermediate formation of cuprous oxide in one case and not in the other. But an intermediate stage of this kind should appear from the curves showing the relation between the amount of chemical change and time whereas no such indication is present. During reduction the copper is on the outside of the granules and the oxide on the inside, while during oxidation the positions are reversed. This might be the cause of the difference, the absence of diffraction colours during reduction being due to the great scattering powers of the metal, which enables a small particle thinly covered with copper to behave as though it were copper throughout.

A more probable explanation seems to be that, while in oxidation each existing spherule of copper is gradually converted into oxide, reduction necessitates the growth of fresh granules, which grow about nuclei of copper, so that the film does not consist of composite units, but is an assemblage of growing granules of copper and diminishing granules of oxide. The kinetics of the reduction lend some support to this suggestion, which is further strengthened by a rather curious phenomenon sometimes observed. If the film is not oxidised beyond about the green stage, then it may sometimes be made to re-traverse during reduction the same colour sequence as it passed through during oxidation, the considerable core of unoxidised copper in the granules being sufficient to maintain the original structure rather than allow the growth of fresh granules.

It should be mentioned that heating films of given colour in a vacuum over several hundred degrees produces no appreciable change in their colour.

Kinetics of the Oxidation and Reduction.

Experiments at high pressures have been made by Joannis,[*] Berger[†] and Stahl.[‡] Berger finds the velocity of oxidation to be nearly independent of the pressure of oxygen for pressures from 1000 to 100 mm., and only to fall off at pressures less than 100 mm.

In these experiments the pressures were low, and the velocity both of oxidation and reduction was dependent on the pressure.

Most attention was given to the gradual activation of the surface. Only typical results will be given in full, and the general nature of the others described.

Oxidation.—The formation of an active granular film is a gradual process requiring a number of oxidations and reductions. A limiting velocity appears ultimately to be reached, provided that the oxidations are confined to the film, and not prolonged so as to eat further into the massive metal below.

The equation connecting oxygen absorption with time is probably very complex, but since the main governing factor is the velocity of diffusion of oxygen through the oxide already formed, an approximate description of the process should be given by the equation

---

[*] 'C. R.,' vol. 158, p. 1801 (1914).

[†] 'C. R.,' vol. 158, pp. 1502 and 1798 (1914).

[‡] 'J. S. C. I.,' vol. 33, p. 1158 (1914).

$$-dp/dt = \frac{kAp}{p_0 - p},$$

where A = total area of all the granules,
p = pressure of oxygen,
p_0 = initial pressure,

thus $p_0 - p$ is proportional to the amount of oxide already formed.

Hence $$kA = 1/t\{p_0 \log p_0/p - (p_0 - p)\}.$$

kA is found not to be really constant, though it should be pointed out that small differences in p make large changes in kA. Nevertheless, the variation in kA over a series of experiments due to the real change in A accompanying the change in structure of the film is large compared with the variations due to the imperfection of the formula. Hence kA serves as an approximate measure of the activity of the film.

1. Oxidation of Fresh Bright Foil (3·19 sq. cm.)
Temperature 302° C.
p_0 = 4·75 mm.

t (minutes).	$p_0 - p$.	kA.
0	0	
30	0·32	
230	0·82	
510	1·56	$6·4 \times 10^{-4}$
750	2·00	$8·0 \times 10^{-4}$
1660	2·93	$9·9 \times 10^{-4}$

2. Oxidation of Reduced Product of 1.
Temperature 302° C.
p_0 = 5·01 mm.

t.	$p_0 - p$.	kA.
0	0	
20	1·38	110×10^{-4}
70	2·15	94×10^{-4}
110	2·31	72×10^{-4}

3. After four Oxidations and reductions.
Temperature 302° C.
p_0 = 3·80 mm.

t.	$p_0 - p$.	kA.
0	0	
7	0·66	
13	1·24	200×10^{-4}
40	1·74	150×10^{-4}

4. Oxidation of Reduced Product of 3.
Temperature 302° C.
p_0 = 4·49 mm.

t.	$p_0 - p$.	kA.
0	0	
5	0·83	180×10^{-4}
10	1·22	200×10^{-4}

5. Oxidation of Reduced Product of 4.
Temperature 302° C.
p_0 = 3·25 mm.

t.	$p_0 - p$.	kA.
0	0	
4	0·28	
20	0·99	100×10^{-4}

In another series with different foil and at 233° C. the fresh foil gave the following kA values $4·2 \times 10^{-5}$, 7·8, 8·3, $8·4 \times 10^{-5}$, its area being 8·94 sq. cm. After a good surface film had been formed by heating in an electrically heated tube to about 360° and oxidising and reducing several times it gave a kA value of about 150×10^{-5}. Although the permanent colour sequence was established it was by no means fully activated for the lower temperature, and after five oxidations at 233°, in none of which the black stage was reached and which were all short enough to ensure that an inappreciable amount of the massive copper underneath had been attacked, kA value of 4800×10^{-5} was reached. After this the activity decreased again.

The values for the experiment in which the greatest activity was exhibited may be quoted as typical:—

t. minutes.	p. obsd. (mm.).	p. calc. for $kA = 0{\cdot}048$.
0	1·23	1·23
2	0·97	0·93
5	0·62	0·61
10	0·39	0·43
15	0·30	0·33
75	0·12	0·08

Although it might have been expected that a film formed by repeated oxidation and reduction at a high temperature would show no further increase in activity in a series of experiments at a lower temperature, this appears not to be true. Fig. 2 illustrates a series made at 233° with a film formed by over forty oxidations and reductions at 400°. During formation oxidation was carried to the black stage each time, during the experiments only to the purple stage.

The curves are numbered in the order in which the experiments were performed. Here, again, the activity reaches a limit and declines slightly. The maximum activity corresponded with a kA value of about 7000×10^{-5}. The limiting rate would naturally be reached when the granular film had attained a condition of maximum stability where it was incapable of further sub-division, but the possibility of loss of activity by poisoning with traces of impurity in the oxygen or hydrogen must not be lost sight of.

Reduction.—Joannis at 100–600 mm. found that the reduction was not retarded by the copper formed to the same extent as the oxidation was retarded by the oxide. Berger found an induction period in the reduction of the oxide. This could not have been due to the necessity for water-vapour to be produced since water-vapour actually retards the reduction.

Both of these observations are confirmed by the low-pressure experiments, but it was found that there was no induction period at first. As the series proceeded and the foil grew more active both with respect to oxidation and reduction an induction period made its appearance.

These observations all fit in with the suggestion that reduction does not proceed through the original granules of oxide, but that fresh nuclei of copper grow at the expense of the oxide nuclei. Any change which depends upon nucleus formation may show an induction period, and often is accelerated in an autocatalytic manner. There are evident signs of such acceleration in the curves showing the progress of the reduction with time.

The growth of fresh granules of copper during reduction offers a simple picture of how the activation by alternate oxidation and reduction comes about.

Order of Magnitude of the Granules Composing the Film.—On the simple assumption already made we have

$$A = \frac{1}{kt}\{p_0 \ln p_0/p - (p_0 - p)\}. \tag{1}$$

If the film consist of n granules of radius r then

$$4\pi r^2 n = A, \tag{2}$$

also m, the mass of oxide in the oxidised film, will be given by

$$4/3\pi r^3 n \times 5{\cdot}9 = m, \tag{3}$$

assuming the normal density (5·9) of the oxide within the limits of the granule.

These determine r if k is known. This is the main difficulty. We will first find some values of r/k. As in most cases the films were not oxidised to the black stage, the amount of oxygen corresponding to light green was used for the calculation of m with the aid of the ratios given in the section on the colours. Three experiments at 233° lead to values of

$$r/k \text{ of } 4 \times 10^{-2},\ 2 \times 10^{-2}, \text{ and } 1{\cdot}5 \times 10^{-2}.$$

At 302° with different copper with films more or less in the state of maximum activity 27×10^{-2} and 29×10^{-2} were found.

To find k we must obtain the value of kA in an instance where A is known. The simplest way is to assume that with new foil when the metal is still in the compact form the measured area of the foil is also that governing diffusion. On this basis a value of k of the order 10^{-5} was found for the 233° series, which leads to a value for the diameter of the granules of the order 0·005 μ. In the series at 302° the value of k arrived at on the same assumption was of the order 3×10^{-4} which leads to a value for the diameter of about 0·17 μ. Owing to the crudeness of the assumption which must be made in finding k these figures only show the order of magnitude, but would appear to represent major limits, since the foil, even in the compact state, almost certainly offers to the attack of the oxygen an area considerably greater than its apparent area. We cannot, therefore, attach importance to the difference in the values found at 233° and 302°, but both sets of values show the degree of dispersion of the metal in the film to be very great. The particles are small enough therefore to assume forms in which the influence of surface tension predominates.

I hope to improve the method of investigation by employing films of the attackable metals deposited on an inert metal.

On Measurements of Electrode Potential Drop with Direct Current and Alternating Current Electrolysis.

By S. Marsh, B.Sc., Ph.D., and A. E. Evans, B.Sc.

(Communicated by E. H. Griffiths, F.R.S. Received April 29, 1922.)

§1.

The investigations to be described in this paper are the outcome of some work by one of the authors on alternating current electrolysis,[*] in which the behaviour of platinum, gold and nickel electrodes, using dilute sulphuric acid and barium hydrate solution as electrolytes, was examined for alternating current with frequencies lying between 25 cycles and 80 cycles per second. Some provisional measurements of the electrode drop were made at that time, but it was felt desirable to examine this feature in greater detail.

The apparatus employed for the purpose is shown in fig. 1. The electrolyte was contained in a glass tube EF, 60 cm. long and 4 cm. diameter, provided with a feed tube D. The ends were closed with rubber stoppers, each pierced with two holes, through which passed tubes holding electrodes A and C and an "explorer" B, and a thermometer T. The electrode C was of platinum, and remained unchanged throughout all the experiments.

A was the experimental electrode, rectangular in shape, area approximately 0·4 sq. cm. Electrodes of double and half this area were also employed to investigate the effect of current density.

It was attached to a short platinum wire which was sealed through the tube, and electrical contact was made by mercury with wires from the supply. The "explorer" B consisted of a fine platinum wire (0·05 mm. diameter) which projected about 1 or 2 mm. from a fine glass tube sealed at B. The distance AB was approximately 1 mm. in all cases. Contact was made with mercury as before.

---

[*] S. Marsh, 'Roy. Soc. Proc.,' A, vol. 97, pp. 124–144.

Alternating current was supplied from a small Westinghouse motor-generator set up to 40 cycles per second, and from a Crompton machine for frequencies up to 80 cycles per second.

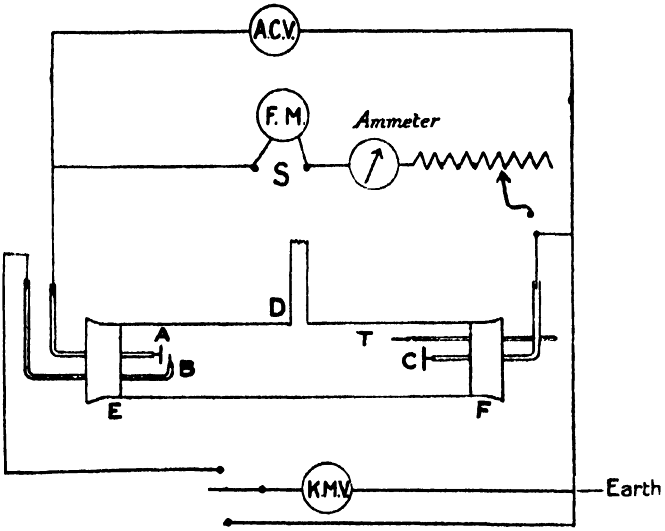


Fig. 1.

The arrangements for the measurement of electrode drop are also shown in fig. 1. The difference in potential between A and B is too small (at any rate when alternating current is used) to allow of direct measurement. It was necessary, therefore, to make simultaneous readings of the potential difference between A and C, and B and C. That between A and C was determined by a Weston standard alternating (and direct) current voltmeter reading to 0·01 volt.

The potential drop between B and C was obtained by a Kelvin multi-cellular electrostatic voltmeter. As the scale of the latter commenced at 40 volts, while the potential drop from B to C was of the order of 10 volts, a small mirror was fixed to a new and more sensitive suspension, and readings made by lamp and scale. The Kelvin instrument was calibrated both with direct and alternating current, and the results were in excellent agreement.

Experiment showed that a current of 0·4 ampère gave readings on the open part of the scale of the multicellular instrument, and, as a consequence, this value of the current was employed throughout. The current density in most cases was approximately 0·5 ampère per square centimetre, counting both faces of the electrode.

The multicellular voltmeter was much more sluggish in its movements than the Weston voltmeter, and in order to reduce any error arising from this cause, the current through the tube was carefully maintained at 0·4 ampère throughout the experiment. The frequency was kept at a constant value as registered on a Hartmann-Braun frequency-meter. Readings were taken every minute on the voltmeters.

A direct determination of the potential drop from A to B is desirable, but its magnitude (usually less than 1 volt) is too small for detection by ordinary methods when alternating current is employed.

Calculation has shown that it would be possible to measure potential differences of that order by means of an electrostatic oscillograph in which low velocity cathode rays produced from a Wehnelt cathode are deflected in passing between condenser plates connected to A and B, and it is hoped to proceed further with the construction of such an instrument.

Normal sulphuric acid was employed throughout as the electrolyte, fresh acid being used for each experiment, while the electrodes consisted of polished platinum, platinum black, gold and nickel.

§ 2.

Summary of Results with Direct Current.

Electrode Acting as Cathode.

The relation of electrode drop to time is indicated in fig. 2, where typical curves are represented for the various metals. Too much importance must not be paid to the actual magnitudes involved, since repetition of an experiment under identical circumstances did not lead, as a rule, to the same quantitative results. The *type* of curve was, however, the same in repetition experiments.

It will be noticed that with platinum, gold and nickel electrodes, the cathode drop increases with time, the curves, especially those for polished platinum, being similar in type to the saturation curves in radio-activity. With platinum, the electrode drop becomes approximately constant after 30 minutes. In the case of gold, the increase is not so much as with platinum, and in one case there is a suggestion of a second rise after about 60 minutes.

With platinum black electrodes, the increase in the electrode drop is more gradual. The evidence, in the case of nickel electrodes, is somewhat conflicting, but the curves on the whole resemble those for platinum and gold.

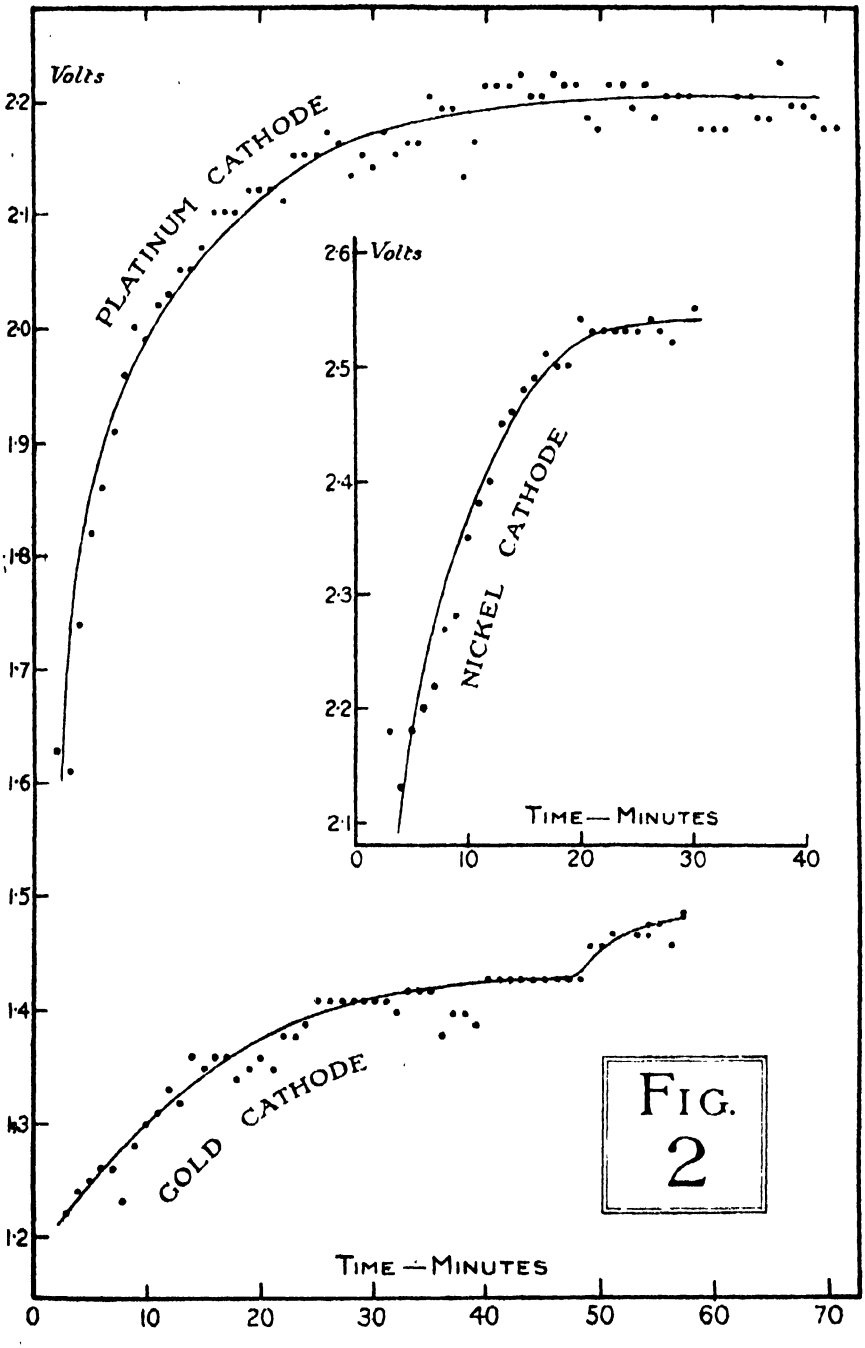

Volts
2·2
2·1
2·0
1·9
1·8
1·7
1·6
1·5
1·4
1·3
1·2
PLATINUM CATHODE
GOLD CATHODE
TIME — MINUTES
0
10
20
30
40
50
60
70
2·6
Volts
2·5
2·4
2·3
2·2
2·1
NICKEL CATHODE
TIME — MINUTES
0
10
20
30
40

FIG. 2

Electrode Acting as Anode.

Typical curves are shown in fig. 3. With polished platinum, the electrode drop decreases until about the 40th minute and then rises. The rates of fall and rise are approximately the same, as are also the amounts of ascent and descent.

In the case of gold, the rate of descent is slightly more rapid and the amount of ascent smaller than the total fall. With platinum black, the variations of electrode drop are not so marked.

§ 3.

Results with Alternating Currents.

As will be seen from fig. 4, the variation of electrode drop with time is very irregular, and there is considerable divergence from the smooth curve drawn to lie evenly between the points. It is, of course, obvious that in this case the ordinates do not measure electrode fall in the same sense as with direct current, for the electrode is alternately anode and cathode; and furthermore, the ordinates represent the differences between simultaneous readings on the alternating current voltmeter across AC and the Kelvin instrument across BC.

Hence the importance of setting up, if possible, a low voltage electrostatic oscillograph, and reading directly the drop between A and B. This voltage would furnish, in addition, a knowledge of the *form* of the wave curve.

Platinum.—For the frequencies examined (25 to 80) the electrode drop decreases during an interval—the length of which depends upon the frequency of the alternating current—and thereafter increases slightly. This is summarised in the following table:—

Size of electrode.	Frequency.	Time to reach minimum electrode drop.
Full	25 cycles.	40 mins.
"	33·5 "	30 "
"	45 "	25 "
"	60 "	30 "
"	80 "	15 "
Half	25 "	100 "
"	33·5 "	60 "
Double	25 "	40 "
"	33·5 "	27·5 "

Gold.—There is the same general tendency as with platinum except that the minimum is reached somewhat earlier.

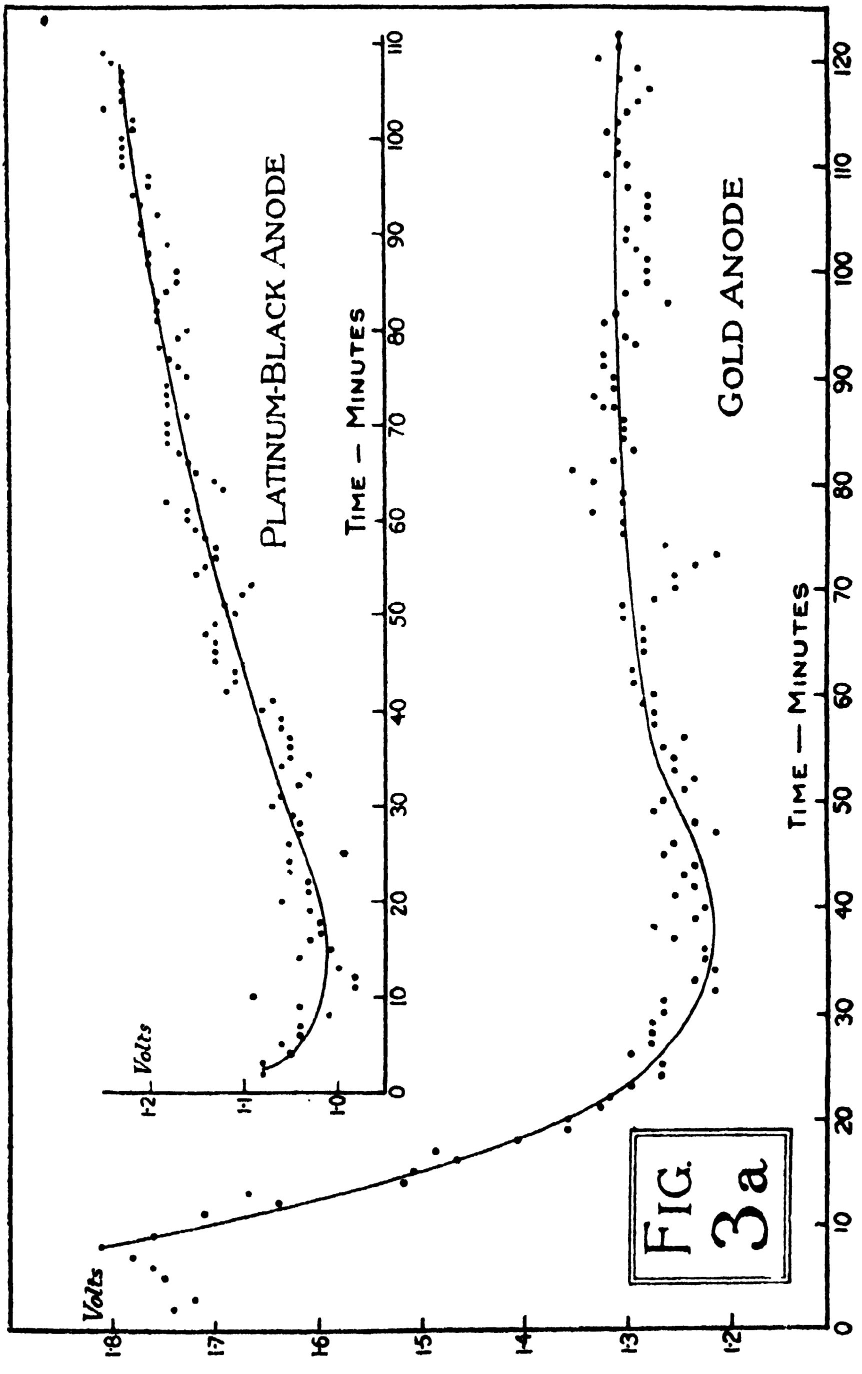

PLATINUM-BLACK ANODE
GOLD ANODE
TIME — MINUTES
TIME — MINUTES
Volts
Volts
1·8
1·7
1·6
1·5
1·4
1·3
1·2
1·2
1·1
1·0
0
10
20
30
40
50
60
70
80
90
100
110
120

FIG. 3a

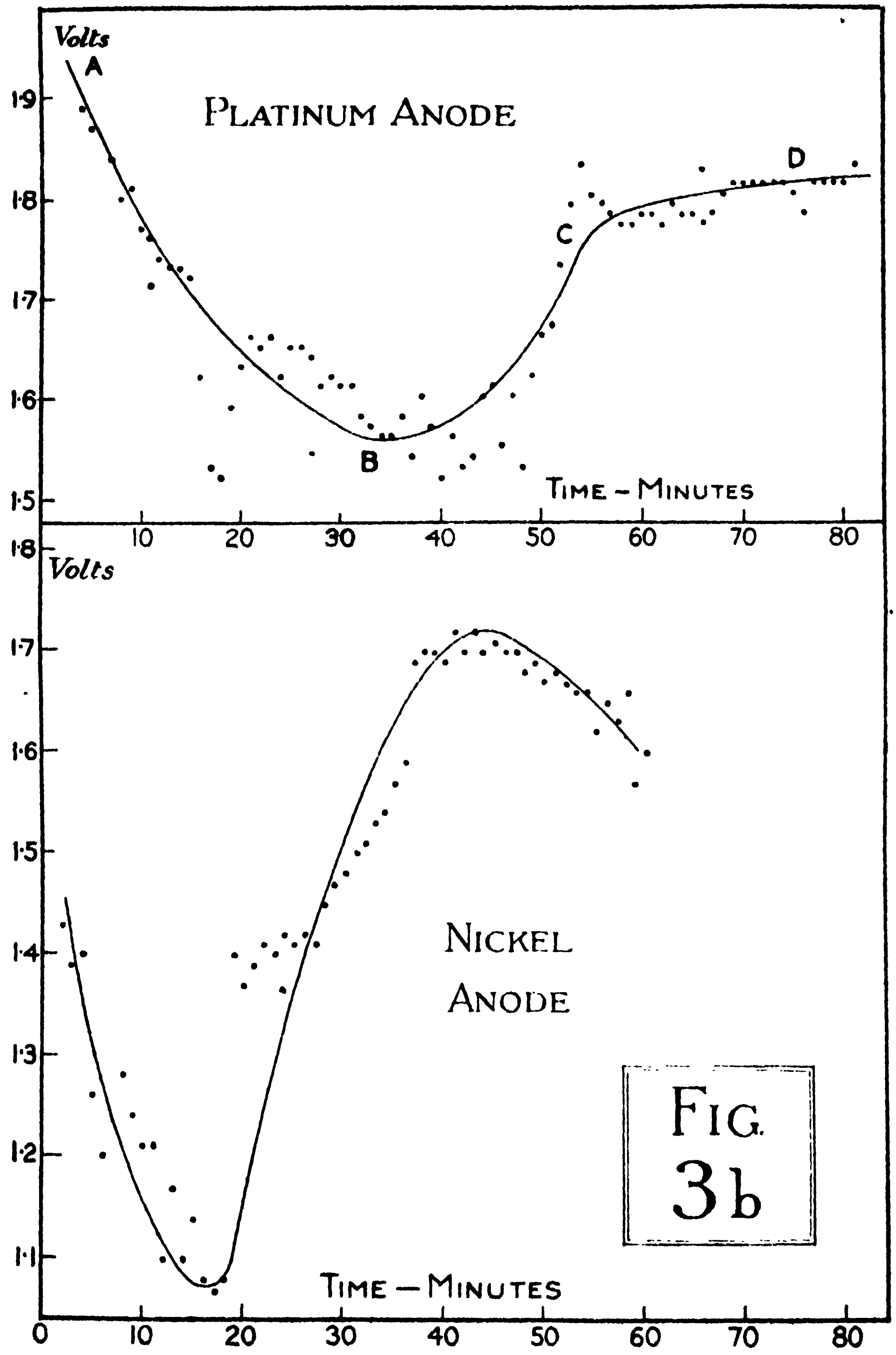

Volts
A
Platinum Anode
B
C
D
1·9
1·8
1·7
1·6
1·5
Time – Minutes
10
20
30
40
50
60
70
80
1·8
Volts
1·7
1·6
1·5
1·4
1·3
1·2
1·1
Nickel Anode
Time – Minutes
0
10
20
30
40
50
60
70
80

Fig. 3b

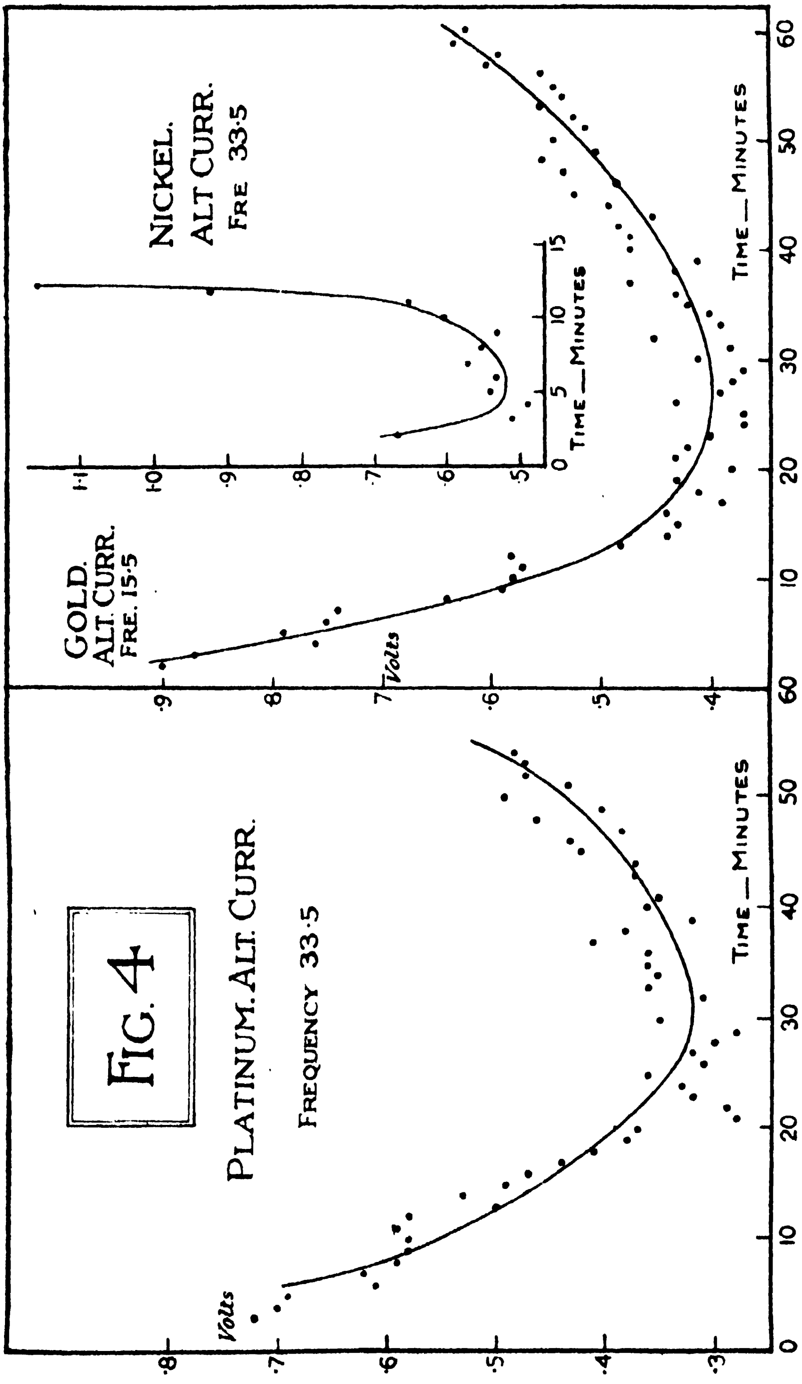


FIG. 4

Platinum. Alt. Curr.
Frequency 33·5
Gold. Alt. Curr. Fre. 15·5
Nickel. Alt Curr. Fre 33·5
Volts
Volts
Time — Minutes
Time — Minutes
Time — Minutes
·8
·7
·6
·5
·4
·3
·9
·8
7
·6
·5
·4
1·1
1·0
·9
·8
·7
·6
·5
0
10
20
30
40
50
60
10
20
30
40
50
60
0
5
10
15

Nickel.—The decrease and subsequent rise is much more rapid than with gold and platinum. The very sudden ascent is due to the solution of the electrode in the electrolyte and the consequent increase in current density.

Black Platinum.—In the majority of cases the readings, though very irregular, suggest a more or less steady voltage.

§ 4.

Discussion of Results.

The phenomena occurring with direct current electrolysis will first be examined. As the measurements are made during the actual passage of the current, the results are not to be confused with the measurements of the ordinary contact difference of potential between metal and acid.

The phenomena taking place at an electrode comprise :—

(1) The evolution of gas, which coats the surface of the electrode with a film of bubbles ;

(2) Absorption and adsorption of gas by the electrode ;

(3) Oxidation of the electrode (in the case of an anode) or formation of hydrides (with cathode) ;

(4) Chemical changes in the electrolyte.

The effect due to (1) above will be, at any rate with direct current, of a fairly constant nature.

Electrode Acting as Cathode.

The gas entering the electrode, being in the ionic state, has a tendency to re-enter the electrolyte, and so sets up a back e.m.f., whose magnitude depends upon the quantity of gas occluded. This quantity will increase with time, the relation between amount occluded and time being represented approximately by a curve of the "saturation" type. It is reasonable to assume that the back e.m.f. will vary in a similar manner, and hence the curves representing the relation of electrode drop to time will be of the type shown in fig. 2.

The more quickly the surface layers of the electrode become saturated with the gas the more rapid will be the consequent increase in electrode drop. With a prepared surface of great area, such as platinum black, it would be possible for occlusion to go on for a considerable period before the metal gets saturated. Hence, we should anticipate a lower initial electrode drop and a smaller rate of increase with time.

It is obvious that any agency which tends to break up the surface, thus exposing a greater area to the action of the gas, will facilitate occlusion, and it is highly probable that the disintegration of the surface, which takes place with alternating current electrolysis, has a very important bearing on the electrode drop in such cases.

Since more gas is occluded by platinum than gold, we should expect that the rise in the electrode drop would be greater in the case of platinum than with gold. This is confirmed by experiment.

Electrode Acting as Anode.

In this case oxidation and occlusion are going on simultaneously, but for the sake of simplicity, we consider these processes separately. We conceive, then, that oxidation of the surface layer of the metal takes place, and that in consequence of the disintegration of surface thereby produced a larger area of metal surface is exposed to the action of the gas, thus allowing oxidation to proceed at an increasing rate.

The surface XY of pure metal (fig. 5), exposed to the action of the gas, would gradually increase up to a maximum; as a consequence, the rate of removal of ionised gas at the metal face, by combination, would also increase to a maximum, and the electrode drop, on this account, would fall to a minimum value.

As the thickness of the oxidised layer increases, however, the amount of ionised gas reaching the virgin surface XY per second becomes smaller, owing to absorption in the layer of oxide to the left of XY. This gradual accumulation of gas in the oxide will produce an increasing back e.m.f. similar in type to that occurring when the electrode acts as cathode.

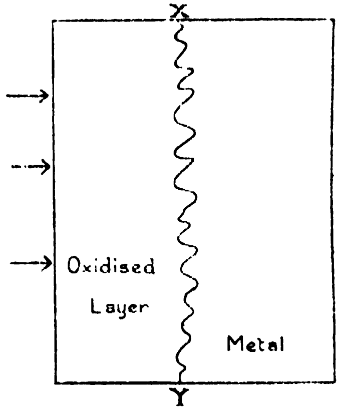


FIG. 5.

Thus in fig. 3, the portion AB of the potential-time curve is attributed to the effect of oxidation, while BCD represents the state of things when the removal of gas by oxidation is more than nullified by the gradual accumulation of gas in the surface layers.

If the film of oxide became detached mechanically or dissolved in the electrolyte, then it is conceivable that oxidation at the metal face would gain the upper hand for a time, and the curve repeat itself. There have been indications of such curves with platinum.

With a platinum black electrode, oxidation would proceed immediately at the maximum rate at the prepared surface; in other words, the initial electrode drop would be smaller than with polished platinum. Occlusion would come into action earlier so that we should expect that the portion AB of the curve in fig. 3 would be much less pronounced, and that the increase in the electrode drop would occur earlier. The experimental results lend support to these suggestions.

In the case of gold, the portion AB of the curve is very pronounced, but the portion BC is not nearly so marked as with platinum. This would indicate that occlusion plays a much less prominent part with gold than with platinum.

In the case of nickel, it must be remembered that the oxide formed is soluble in the acid,* and the metallic surface is roughened, exposing consequently a larger area for oxidation. Hence, the initial drop AB in the curve would be expected. The gradual accumulation of gas in the surface layers of the nickel (or of the oxide, if the metal is oxidised at a greater rate than it can be dissolved) would give rise to the portion BC of the curve.

§ 5.

Phenomena with Alternating Current.

There is evidence that the minimum electrode drop is reached sooner with high than with low frequencies (see Table, p. 332).

In addition to the processes of oxidation and occlusion already discussed with direct current electrolysis, there is the chemical action between the gases liberated in each half-cycle and those present, either in a state of occlusion in, or in combination with, the electrode.

During the first anodic half-period some oxidation occurs, and, in addition, some oxygen enters the electrode. In the succeeding half-period, the hydrogen liberated reduces the oxide and combines with the oxygen occluded in the oxide; in addition, some hydrogen is retained at the surface, while some diffuses into the deeper layers. The oxygen of the third half-period is able, since the electrode surface is more broken up, to enter more readily. It combines with the hydrogen in the surface layers; oxidises these layers of spongy metal; and, in addition, diffuses into the layers beneath.

Since hydrogen diffuses more quickly than oxygen, the deeper layers gradually accumulate this gas. Hence, there is less tendency for the hydrogen liberated in later cathodic half-cycles to diffuse into these layers, and this hydrogen is confined to the layers nearer the surface, thus facilitating the removal of oxygen of anodic half-cycles.

Thus, as electrolysis proceeds, more and more of the gas produced in any half-period is used up, and finally (subject to conditions stated in the previous paper) evolution of gas ceases.

---

* *Loc. cit.*, pp. 133, 138, § 7.

So far the action of the electrolyte has been neglected. During the anodic half-cycle, persulphuric acid is formed. Hence there is less oxygen available for oxidation and occlusion. Consequently, a surplus of hydrogen is left over after each cycle, so that even after gas evolution has ceased a gradual accumulation of hydrogen still goes on.

We should expect the electrode drop to follow the same sequence of changes as with the anodic fall with direct current, viz., it should decrease as more and more of the gas is removed by chemical action, and ultimately a slight rise should set in owing to the gradual accumulation of hydrogen in the electrode.

These effects have been observed in the great majority of cases with alternating current. Furthermore, it is interesting to record that the electrode drop reaches its minimum value at about the same time that the gases cease to be evolved. As the frequency increases, the more frequent oxidation and reduction will break up the surface layers more effectively, thus facilitating the recombination of the gases; hence, the minimum electrode drop will be reached earlier.

Variation of Current Density.

With electrodes of increasing area (in other words, with decreasing current density), one would expect the surface effect (adsorption) to be more, and the diffusion effect to be less pronounced. Hence there is from the start a more effective recombination (or action) between the gases liberated in successive half-cycles, and we should expect the minimum electrode drop to be reached more quickly. (The subsequent rise also should be smaller, since, with a more pronounced surface effect, there will be less accumulation of hydrogen in the interior. There is ample confirmation of the first conclusion (see Table, p. 332), but the evidence for the second is not sufficiently definite.)

Effect of Different Metals.

The theory outlined above represents the general behaviour of all the metals examined. Taking polished platinum as typical, the following minor differences will be noticed between it and the other metals:—

Gold.—The results with direct current indicate that the effect of oxidation is more pronounced than that of occlusion. The surface effect referred to above should have greater influence than diffusion, and the minimum electrode fall should be reached more quickly than with platinum, and the subsequent rise should be smaller.

Nickel.—Solution occurs, the rate of solution increasing with frequency. Solution of the electrode is equivalent to a rapid increase in the current density, and we should expect the diffusion effect to be pronounced. The electrode would get saturated with hydrogen and would offer a big back e.m.f.

Experiment shows that the initial drop is followed by a very sudden rise as the electrode dissolves.

Platinum-Black.—Owing to the prepared surface, the rate of recombination should be very great from the commencement. No gases, in fact, are evolved with alternating current.

Experiment shows that the electrode fall is more or less steady.

Summary.

The variation with time of the electrode potential drop of various metals has been examined both with direct and alternating current. An attempt has been made to interpret these variations as being due to oxidation and occlusion phenomena within the electrodes. The conclusions confirm the results of the previous investigation on the volume of gases evolved during alternating current electrolysis.

We desire to express our thanks to Dr. E. H. Griffiths, F.R.S., for his interest in the paper, and to Mr. W. H. Sewell for his assistance in taking the observations.

Variation of the Intensity of Reflected X-Radiation with the Temperature of the Crystal.

By IVOR BACKHURST, B.Sc.

(Communicated by Sir William Bragg, K.B.E., F.R.S. Received August 29, 1922.)

On account of the increasing importance of the temperature factor in the theory of the analysis of crystal structure by means of X-ray reflection, and in view of the fact that no experimental work on the temperature effect has been published since the original preliminary investigation by Prof. W. H. Bragg[*] in 1914, it is hoped that the present account will be of interest, in which the work of Bragg is extended to a variety of crystals and continuous temperature curves have been obtained ranging up to 950° C.

---

[*] See "X-Rays and Crystal Structure," by Sir William Bragg and W. L. Bragg, also 'Phil. Mag.,' May, 1914.

Theories of the effect of temperature on X-ray reflection have been given by C. G. Darwin* and P. Debye,† and these agree in that the intensity of reflection—

(1) Decreases as the temperature of the crystal increases;

(2) Decreases more rapidly as the angle of reflection increases.

These conclusions, which can be arrived at by general considerations if it is assumed that the diffracting centres are in motion due to heat, are supported by the experiments of Bragg, and also by those here described, in which, with the notable exception of the sapphire and ruby, the intensity of higher order spectra has been found to diminish more rapidly with rise of temperature than that of lower order.

The method of experiment adopted by Bragg was to place the crystal inside an electric heater, furnished with mica windows suitably placed to allow of the ingress and egress of the X-rays, and provided with a thermometer arranged so that its bulb was immediately above the crystal. Intensities of the several orders of the reflection from different planes of rock-salt and sylvine were compared at room temperature and at 370° C. and 311° C. respectively.

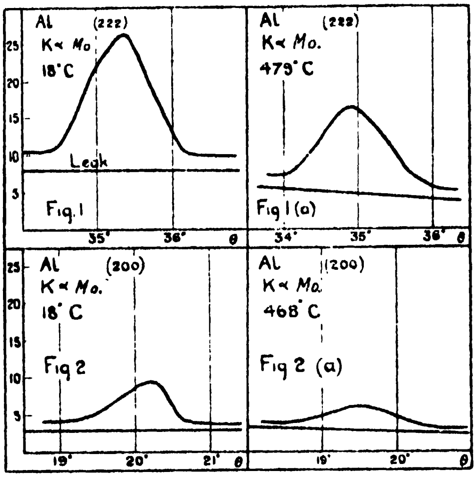
In the present research a similar method was employed, the electric heater used being constructed to admit of a higher temperature range and the introduction of a nitrogen atmosphere, while a thermocouple was used to indicate the temperature of the crystal. It was also found necessary to use a special type of mounting for the crystal, in order to prevent rapid loss of heat by conduction, and at the same time allow of freedom of adjustment. Experiments were made with a number of crystals, and among these, aluminium, carborundum, graphite, diamond, sapphire and ruby furnished definite results which, while of the same order as predicted by theory, do not show good numerical agreement with the latter. This general result is the same as that arrived at by Prof. W. H. Bragg.

Aluminium.—In figs. 1 and 2 are shown the curves obtained for aluminium using molybdenum K-radiation, the width of the bulb, crystal and ionisation chamber slits being in each case 2 mm. The crystal reflection was integrated at a uniform rate over the whole angular range for which reflection occurred, half a degree being found sufficient in each case. The leak shown represents the deflection obtained when the ionisation chamber was just out of alignment. A considerable number of crystals were tried, many being unsatisfactory, as they gave progressively smaller reflections on repeated heatings, indicating a breaking down of the one complete crystal into smaller ones differently oriented.

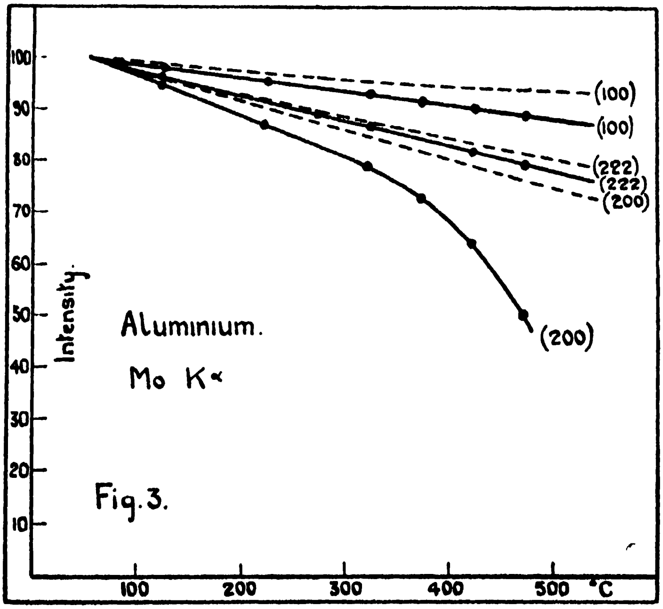
---

* 'Phil. Mag.,' February and April, 1914.

† 'Annalen der Phys.,' vol. 43, p. 87 (1914).



In fig. 3 the intensity for the reflection in question is taken as being 100 at 50° C., the ordinates to the continuous curves representing the corre-



sponding intensities at higher temperatures. The points are each obtained by taking the mean of a series of observations, first for 50° C., then for the temperature indicated, then for 50° C., and so on; only the results for crystals giving consistent series of ratios being retained. The ionisation chamber was set to take in the whole beam, and the reflection from the crystal integrated as in the original method. The deflection so obtained was taken to represent the combined intensity of the characteristic beam and the heterogeneous radiation. The intensity of the heterogeneous radiation was taken to be the same as that in an equal angular range just beyond the range of the characteristic radiation. To obtain this measure of the heterogeneous radiation, the crystal stage and ionisation chamber were moved to the proper positions and the reflection integrated as before, the difference between the two deflections being taken as proportional to the intensity of the characteristic radiation.

Owing to the expansion of the crystal, the position of the ionisation chamber had to be varied with the temperature.

Values obtained are given below :—

Temperature (t° C.).	Intensity (I).	Alteration of glancing angle $\theta/2$.
	(100) plane $\theta/2 = 10^\circ\ 4'$.	
50	100·0	0′
118	98·1	1′
218	96·0	3′
318	92·8	4′
368	91·0	5′
418	90·5	5′
468	88·9	6′
	(222) plane $\theta/2 = 17^\circ\ 40'$.	
118	96·6	3′
268	88·8	7′
318	96·7	8′
418	82·6	11′
468	79·9	13′
	(200) plane $\theta/2 = 20^\circ\ 28'$.	
118	95·3	3′
218	87·4	7′
318	80·1	10′
368	73·8	11
418	64·0	13′
468	50·8	15′

The expression obtained by P. Debye* for the temperature variation of intensity in the case of a crystal in which the atoms are in simple cubic array is of the form

$$(1)\ \mathrm{I} \propto e^{-\mathrm{P}(1-\cos\theta)\phi(x)/x}, \qquad \text{or } (2)\ \mathrm{I} \propto e^{-\mathrm{P}(1-\cos\theta)(\frac{1}{4}+\phi(x)/x)},$$

where $\mathrm{P} = \dfrac{0{\cdot}571 \times 10^{-12}}{\mathrm{A}\,.\,\mathrm{T}_c\,.\,\lambda^2}$, $x = \dfrac{\mathrm{T}_c}{\mathrm{T}}$, $\phi(x)$ is a calculated function of x, given by

$$\phi(x) = \frac{1}{x}\int_0^x \frac{\xi d\xi}{e^\xi - 1}.$$

A = atomic weight, T is the absolute temperature, and T_c is a temperature peculiar to the substance under consideration, which Debye calls the characteristic temperature.

---

* 'Annalen der Phys.,' vol. 43, p. 87 (1914).

In either (1) or (2) we have, if I_1, x_1, T_1, and I_2, x_2, T_2, are corresponding values, that

$$\frac{I_2}{I_1} = \frac{e^{-P(1-\cos\theta)\phi(x_2)/x_2}}{e^{-P(1-\cos\theta)\phi(x_1)/x_1}} = e^{-P(1-\cos\theta)[\phi(x_2)/x_2-\phi(x_1)/x_1]},$$

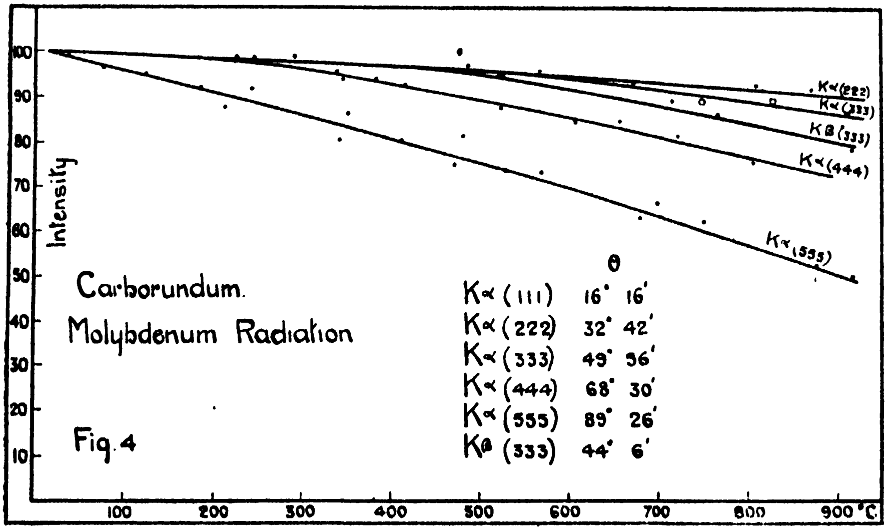
or since $n\lambda = 2d\sin\theta/2$

$$\text{(3)} \qquad \frac{I_2}{I_1} = e^{-0{\cdot}571\times10^{-12}\,.\,n^2\,.\,[\phi(x_2)/x_2-\phi(x_1)/x_1]/[A\,.\,T_c\,.\,2d^2]}.$$

For aluminium A = 27·1, and Debye* gives the value of T_c as 396.

For the (100) planes the spacing $d = 2{\cdot}03$ Å, and for the (111) planes $d = 2{\cdot}34$ Å. Putting in these values, the curves shown dotted in fig. 3 were obtained. It has not been thought necessary to apply the small corrections for the lack of complete homogeneity of wave-length or for the finite breadth of the beam. Although the agreement is quite good for the (222) spectra, there is considerable divergence between the theory and experiment in the case of the (200).

Carborundum.—In fig. 4 the experimental curves for Carborundum are



shown. A crystal with a fine (111) face was used, and there was no trace of deterioration even after repeated heatings to over 900° C. Observations were made, first with the temperature rising slowly, then with gradual cooling. The temperature was measured by means of a platinum-iridium thermocouple enclosed in a fine transparent silica tube placed immediately above the crystal. With this arrangement the lag of the couple was found to be very small.

---

* 'Annalen der Phys.,' vol. 39, p. 812 (1912).

The electric heater used for aluminium was unsuitable for temperatures over 600° C., and considerable time was spent in trying various forms of electric furnaces in order to extend the temperature range. The experimental difficulties were peculiar in that the furnace had to be very small and yet had to produce a uniform high temperature in the neighbourhood of a crystal less than 3 cm. from the large brass table of the spectrometer. At the same time, it was necessary to have in the furnace wall opposite the crystal two windows for the ingress and egress of the rays, nothing more than very thin mica being permissible for the windows, on account of the necessity for reducing absorption of the rays to a minimum. In addition to this, the furnace had to be capable of attaining a high temperature quickly, as it is impossible to ensure the constancy of the X-ray emission over a protracted period. Ease of control of the actual temperature was also necessary. Furnaces using mixtures of metallic oxides or graphite as resistances gave high temperatures which were hard to control; also it was difficult to avoid local over-heating. Wire-wound furnaces offered the only practicable solution. Molybdenum winding was not tried, owing to the apparent difficulty of making such a furnace with mica windows sufficiently gas-tight for the use of a hydrogen atmosphere. Nichrome winding was evidently the most suitable, but, on account of the rather low melting-point of this material, a high temperature gradient between the winding and the interior of the furnace had to be avoided, so that a silica core was not permissible. The furnace that was found satisfactory was constructed as follows:—

A tube was made of alundum $1\frac{1}{8}$ inch internal bore, 8 cm. long, and 3 mm. thick: two windows were cut in it for the ingress and egress of the X-rays. The lower end was recessed, and a spiral groove of 1 mm. depth and 2·5 mm. pitch made throughout its length, a hole being bored near either end into which nickel rods could be bolted. The nichrome wire was wound red hot into the spiral and passed at the ends through the nickel rods, the latter keeping the wire in position and serving as terminals. The whole of the nichrome winding, and outside of the tube, was then covered with alundum cement, dried and baked until the cement matured. In this way the whole of the heating element was enclosed (thus reducing risk of oxidation or corrosion) in a chemically very inert substance which, although a good electrical insulator even at high temperatures,* is at the same time a comparatively good heat conductor. Nichrome wire of heavy gauge was used to lead the current to the nickel rods, and in this way corrosion of terminals entirely avoided. The alundum tube was held in position by the three recesses at the lower end, which fitted over narrow steel projecting pieces of the open-work steel base of the outer steel cylinder. Loss of heat by conduction through the frame was thus reduced to a minimum. The space between the core and the outer enclosing cylinder was lagged with magnesium oxide, asbestos being used only where absolutely necessary. Asbestos sheets were used to separate the furnace case and spectrometer table and magnesium oxide to protect the latter from the interior of the furnace. The top of the furnace was closed by four sheets of asbestos board, protected by an asbestos high-temperature compound. With this furnace a current of under 4 ampères (42 watts) produces a rise of temperature of 850° C. in three-quarters of an hour, and over 900° C. in 1 hour.

---

* Electric resistivity of alundum is 1·8 megohm-cms. at 1020° C., and of alundum cement, 6000 ohm-cms. at 1100° C. See E. F. Northrup, 'Met. and Chem. Eng.,' vol. 12, p. 125 (1914), and Hering, 'Met. and Chem. Eng.,' vol. 13, p. 25 (1915).

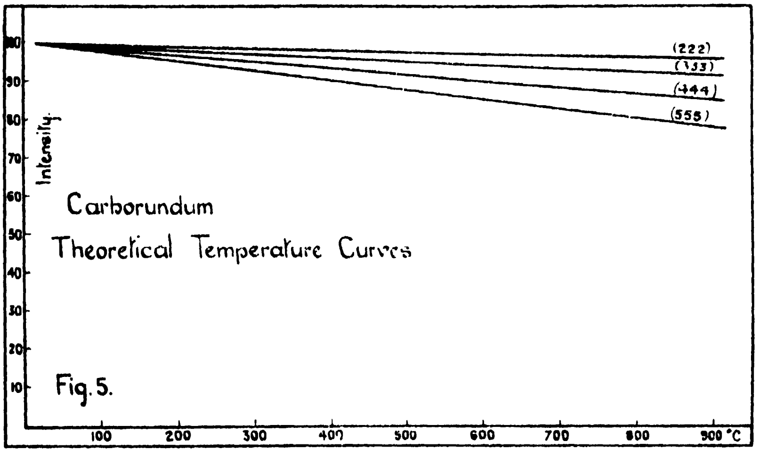
A metallic mounting could not be used for the crystal on account of the rapid conduction of heat through the metal. The mounting adopted consisted of a narrow alundum cylinder surmounted by fire-clay in which the crystal was set.

The points shown in the case of carborundum represent single observations, each the result of rotating the crystal uniformly through 1°, with the ionisation chamber fixed. The alteration of θ with temperature was very small, being only approximately 40 minutes for the Kα (555) spectra at 900° C. Below are given some of the deflections obtained :—

Temperature (t° C.).	Deflection (δ).	Intensity (I).
Kα (222)—		
15	21·2	100·0
292	20·9	98·6
413	20·3	95·8
490	20·7	97·6
565	20·2	95·3
671	19·6	92·5
827	18·9	89·1
872	19·4	91·5
Kα (444)—		
16	23·0	100·0
246	22·7	98·7
337	22·0	95·7
528	20·1	87·4
605	19·3	83·9
719	18·7	81·4
808	17·3	75·0
Kα (555)—		
15	6·1	100·0
213	5·3	86·9
411	4·9	80·4
670	3·8	62·3
912	3·0	49·2

The deflections tabulated are proportional to the intensity of the characteristic radiation only, the deflection for the heterogeneous radiation in each case having been subtracted. In the case of the fourth and fifth orders, however, the amount of the latter was exceedingly small.

It will be seen from the figure that the temperature effect on the intensity becomes rapidly more marked as the value of θ increases; considerably more so than the theory for regular crystals would lead us to expect. If the theory, as it stands, be applied to carborundum, the curves shown in fig. 5 are obtained. They were calculated as follows:—



Carborundum
Theoretical Temperature Curves

Fig. 5.

The specific heat of carborundum at low temperatures has been determined by Paul Günther,* and taking his value for the molecular heat at 73·5° absolute, *i.e.*, 0·442, we have for the ratio

$$\frac{\text{Molecular heat at } 73{\cdot}5}{\text{Maximum molecular heat}} = \frac{0{\cdot}442}{2 \times 5{\cdot}995} = 0{\cdot}03711.$$

From a Table which Debye† gives showing the values of $1/x$ for values of the above ratio, we have:—

$$1/x = 0{\cdot}0775 \text{ and } T_c = \frac{73{\cdot}5}{0{\cdot}0775} = 949.$$

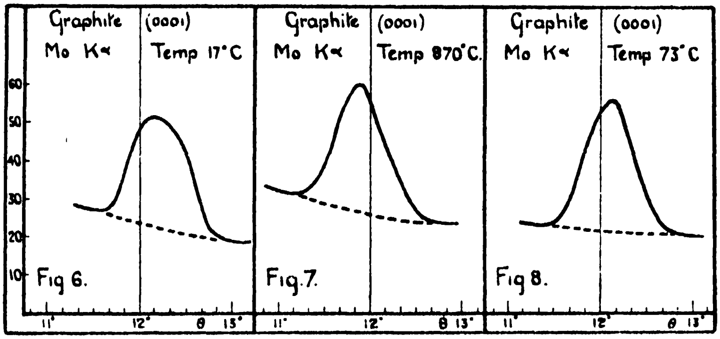
By inserting this value for the characteristic temperature of carborundum in expression (3), together with the values $A = 40{\cdot}3$, $d = 2{\cdot}54$ Å., the curves shown in fig. 5 have been calculated. As will be noticed, it follows from the theory that the curves for the different orders of the same plane should remain unaltered with change of wave-length; for example, Kα (333) should give the same curve as Kβ (333). Referring to fig. 4, it will be seen that a difference exists which, on the whole, seems rather greater than could be accounted for by experimental error.

---

* 'Annalen der Phys.,' vol. 51, p. 840 (1916).

† 'Annalen der Phys.,' vol. 39, p. 803 (1912).

Graphite.—In figs. 6, 7 and 8 the (0001) spectra of graphite for molybdenum Kα-radiation are shown in the order of experiment. Since it was necessary to pass a slow stream of nitrogen through the furnace, in order to prevent oxidation of the graphite, the temperature range was somewhat reduced. It will be noticed that the temperature increase of 850° C. produced an alteration in θ of approximately 17 minutes. This corresponds to an average coefficient of expansion of $26{\cdot}7 \times 10^{-6}$, more than three times the largest value obtained by Fizeau[*] and seven times the largest value obtained by Muraoka.[†] Much more recent determinations for Acheson graphite by Day and Sosman[‡] give $10^6\alpha = 0{\cdot}55 + 0{\cdot}0032\,t$, where $\alpha = [1/\mathrm{L}]\,.\,dl/dt$ between 0° and 1500° C. For the range 0° to 850° C. this gives $[\mathrm{L}-\mathrm{L}_0]/\mathrm{L}_0 = 0{\cdot}001624$ or an average coefficient of expansion of $1{\cdot}91 \times 10^{-6}$, and would produce an alteration in θ of only 1·21 minutes. The observed value is fourteen times as great as this. The X-ray method of measurement gives the expansion along the axis only; other methods give results which are more or less averages of expansion in all directions in the crystal. Hence it would appear that in graphite the whole expansion takes place in the direction perpendicular to the (0001) plane, and it seems probable that a contraction occurs in that plane. It is difficult to test this directly, since planes in graphite perpendicular to the (0001) give spectra with very indefinite maxima and in close proximity to those given by planes that are considerably inclined to the perpendicular. This is unavoidable on account of the fact that the most uniform graphite is far from being a single crystal; the (111) planes of the small crystals, of which the substance is composed, are more or less parallel to each other, but otherwise the arrangement is irregular. There is sufficient evidence, however, to emphasize



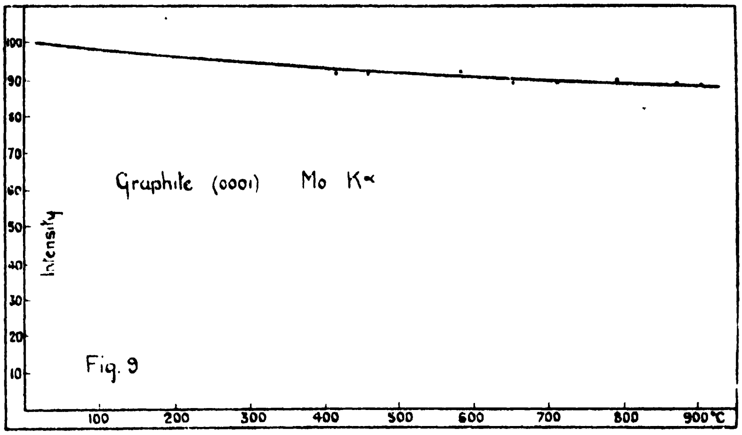
---

[*] 'C. R.,' vol. 68, p. 1125 (1869).

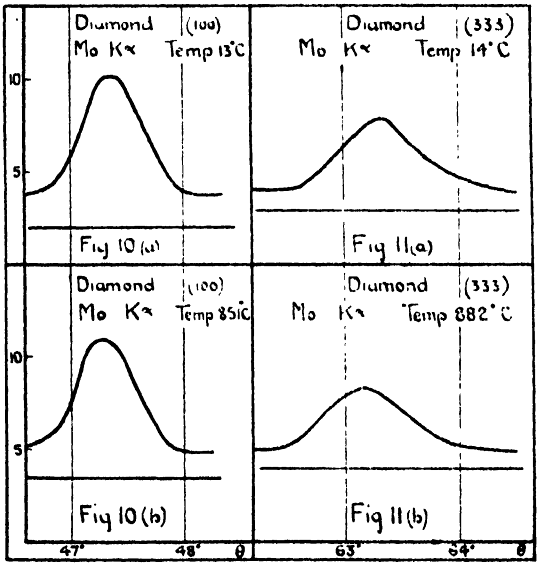
[†] 'Wied. Ann.,' vol. 13, p. 307 (1881).

[‡] 'Washington Acad. Sci. Journ.,' pp. 284–289, June 19, 1912.

the weakness of the bonds perpendicular to the (0001) plane compared with those parallel to it. The intensity curve for the (0001) spectra is shown in fig. 9.



Diamond.—Figs. 10 and 11 are examples of curves obtained with the diamond. The alteration of diffraction angle with expansion corresponds to a



mean coefficient of $2{\cdot}7 \times 10^{-6}$. The expression found by Fizeau* for the range 10° to 90° C. is

$$L = L_0 (1 + \alpha\theta + \beta\theta^2),$$

where $\alpha = 0{\cdot}60 \times 10^{-6}$ and $\beta = 0{\cdot}72 \times 10^{-9}$.

---

* 'C. R.,' vol. 68, p. 1125 (1869).

If Fizeau's formula were assumed to hold over the range of these experiments, viz., 14° to 882° C., the mean coefficient would be $6{\cdot}8 \times 10^{-6}$, much greater than that actually found.

No change in intensity that could be measured with certainty was found for any of the spectra investigated, although in the case of the (333) peak a slight diminution was indicated. The intensity measurements, however, showed that the change was of the order of experimental error, viz., about 2 per cent.

Since the bonds between the carbon atoms in diamond must be exceptionally strong, a very small thermal agitation is to be expected, and fig. 12 shows the small variation of intensity that should exist according to the theory of Debye. Even so, the variation as determined by experiment is considerably smaller.

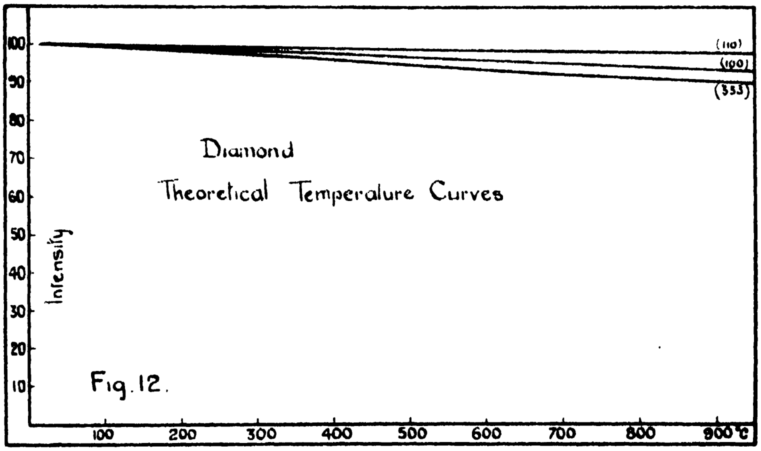


Fig. 12.

The diamonds experimented on were small and completely bathed in the incident beam of rays. The crystal itself was mounted in alundum cement held in a steel cup 4 mm. in diameter, which could be rotated about a vertical axis relative to the vertical segment of a circle to which it was attached, the latter sliding in another circular segment and so capable of rotation about a horizontal axis passing through the centre of the crystal. The whole of the steel mounting was little more than a centimetre high and was cemented to a fireclay cylinder. In this way a very considerable freedom of movement of the crystal was obtained, so that any plane could be brought into position without any part of the mounting interfering with the direct beam. The fireclay foundation was necessary to prevent rapid loss of heat by conduction. A nitrogen atmosphere was used in the furnace.

Ruby and Sapphire.—The curves obtained for the ruby and sapphire are shown in fig. 13, and are of special interest as being the first attempt to examine the behaviour of a crystal which is not cubic, but possesses only the symmetry of the trigonal system. It will be noticed that, while the change in intensity for the (1$\bar{1}$0) spectra is very small, that for the (2$\bar{2}$0) is large, also that the intensity of the (111) spectra falls off more rapidly than the (222). This apparently anomalous behaviour can, however, be explained by assuming that the distance between the centres of the aluminium atom pair remains constant and equal to the diameter of the aluminium atom during the expansion of the crystal. It can be shown that the distance (γ) between the centres of the pair of Al atoms is 2·73 Å.U., since the value for the intensity ratio of the 1st to the 2nd order spectra of the (111) plane, calculated on this assumption, agrees with that experimentally determined at normal temperature. Moreover, the relative intensity of the two orders changes very rapidly as γ changes, enabling the latter to be determined with great precision. This is shown in the following Table*:—

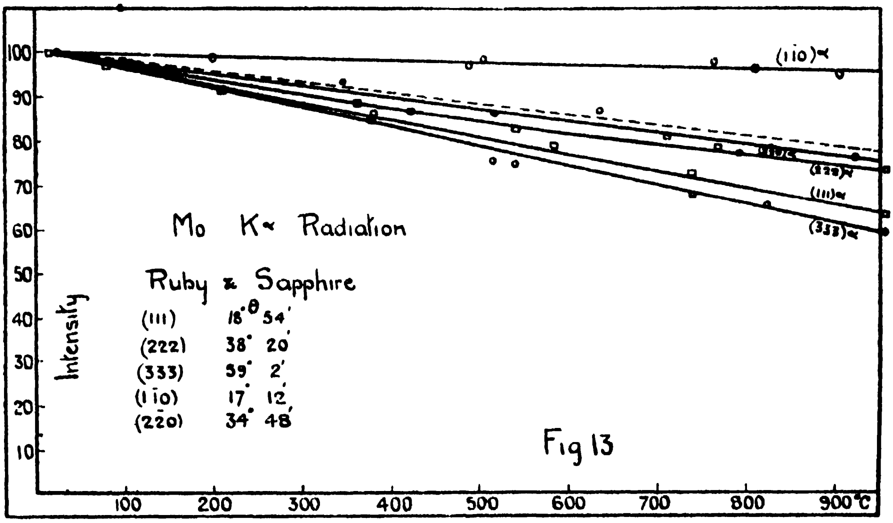


Fig 13

γ	Amplitude on arbitrary scale.		Expected relative intensity.
	1st order.	2nd order.	
2·86	19·8	25·8	1·7
2·76	13·7	39·8	0·47
2·66	8·5	55·0	0·096
2·5	0	79·1	0

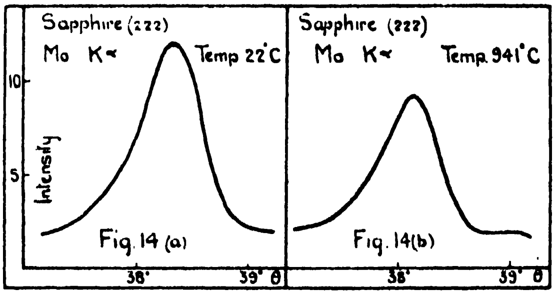
The relative intensity at room temperature is 0·33 (correponding to $\gamma = 2·73$).

---

* I am indebted to Sir William Bragg for these figures.

The decrease in this value observed with rise of temperature could be produced by a decrease in γ, or a proportionate increase in the plane spacing.

The mean coefficient of expansion, as deduced from the observed spectral shift in fig. 14, is $5{\cdot}01 \times 10^{-6}$, and would, therefore, increase the spacing between 15° and 950° C. in the ratio of 1·00469 : 1. Dividing 2·73 by 1·00469 we obtain 2·717, and for this value of γ the relative intensity is 0·27, or 82 per cent. of its previous value. Hence it follows that, if the atom pair remain in contact during the expansion of the crystal lattice, the relative intensity of the 1st and 2nd orders at 950° C. is only 82 per cent. of what it would have been had the distance between the pair centres increased proportionally with the lattice. In the latter case the (111) curve would lie in the normal manner above the (222) curve, and, assuming this curve to be unaltered, would lie in the position shown dotted. The exactness of the coincidence cannot be real, but it is clear that the temperature effect can be explained by natural suppositions concerning the changes which temperature would make in the relative distances and the positions of the atoms of a non-cubic crystal.



In conclusion, I should like to express my thanks to Sir William Bragg for the kind interest he has taken in this work, and to the Department of Scientific and Industrial Research for their grant in aid of the same.

Summary.

In the present investigation Prof. W. H. Bragg's early work has been extended to a variety of crystals and higher temperatures. General agreement only is found with the theories of C. G. Darwin and P. Debye.

Aluminium.—A very marked decrease in intensity was observed with rise of temperature, and fair agreement with P. Debye's theory obtained for the (100) and (222) spectra.

Carborundum.—A special furnace was constructed for temperatures up to 960° C. and no deterioration of the crystal was observed. The decrease in intensity with rise of temperature was much greater for the higher order spectra, and different curves were obtained for the Kα (333) and Kβ (333) spectra.

Graphite.—Only for the cleavage plane reflection was it possible to obtain a definite temperature-intensity curve, and for the direction perpendicular to this plane an unusually high coefficient of expansion was measured.

Diamond.—No decrease in intensity was found that could be measured with certainty, and a very small thermal agitation would be expected on account of the diamond structure's great strength.

Ruby and Sapphire.—An anomalous effect was observed, since the decrease of intensity of the (111) spectra was greater than that of the (222). This may be completely explained by assuming that the atoms of the aluminium pair remain in contact and do not share in the expansion of the lattice.

An Investigation of the Colour Vision of 527 *Students by the Rayleigh Test.*

By R. A. Houstoun, D.Sc., Lecturer on Physical Optics in the University of Glasgow.

(Communicated by Prof. A. Gray, F.R.S. Received June 10, 1922.)

Lord Rayleigh discovered, in 1881,* that if homogeneous yellow is matched with a mixture of homogeneous red and homogeneous green, some persons require much more red, others much more green, in the mixture than the normal. The former have been called red anomalies, the latter green anomalies, and both together have been grouped under the name of anomalous trichromats, because, according to Rayleigh, they are generally not dichromats. Rayleigh obtained matches for 23 male observers. Of these 16 agreed with himself within the errors of observation, and were regarded as normal. Five of the remainder were green anomalies and two red anomalies. Among seven female observers there was not one whose colour vision differed sensibly from his own. Rayleigh did not publish his data, but states that one of the green anomalies had the ratio green to red 2·09 times as great as himself, and one of the red anomalies had the ratio red to green 2·6 times as great as himself. He states that, although the number of observers is insufficient for statistical purposes, it is evident that the peculiarity is by no means rare, at least, among men; he also states that it would seem as if normal colour vision were not of the nature of an average from which small deviations are more probable than larger ones, but that this would require confirmation.

---

* 'Nature,' vol. 25, p. 65 (1881).

In 1890, Schuster* published an account of a more extensive investigation carried out with the same apparatus; 75 individuals were examined; of these three proved colour blind; four, of whom three belonged to the same family, were green anomalies, while one was a red anomaly. Schuster states: (i) that there cannot be any doubt of the real existence of small differences following the ordinary law for deviation from a mean; and (ii) that the larger differences seem certainly more frequent than the distribution of small differences would lead us to expect.

Other investigations on the subject have been carried out by Von Maltzew,† Gutmann,‡ and Edridge-Green.§ Von Maltzew examined 17 cases who were obviously not a random distribution. Gutmann, unfortunately, gives no data, and always takes for granted the existence of two well defined classes of anomalous trichromats, the green and red anomalies, so his work is useless from a statistical point of view. Edridge-Green examined 100 women students, and found that 86 were normal, 10 red anomalies, two green anomalies, while two occupied intermediate positions, but he also does not publish individual readings.

The question of the existence of the green and red anomalies as well defined classes, and their relation to the colour blind, is therefore by no means settled yet, and it was thought desirable, as part of the survey of the colour vision of students which is at present being carried on at Glasgow, to devote special attention to this question. Consequently, during the past year, the colour vision of 527 students has been tested by an apparatus similar to Lord Rayleigh's, and the present paper describes the results.

The apparatus consisted of a spectrometer fitted with a single dense flint prism. The latter was set at minimum deviation for Na light. Between the prism and the telescope object-glass was set up a vertical glass plate, on which was cemented a Wollaston double-refracting prism made in Iceland spar; this glass plate was held in a grating holder. The light which passed through the Wollaston prism was decomposed into two beams polarised respectively in the horizontal and vertical planes; these two beams at the same time suffered deviations, one towards the base of the prism and the other away from it. The light which did not pass through the Wollaston prism, but only through the glass plate to which the latter was cemented, arrived at the object-glass of the telescope without further deviation.

---

* 'Roy. Soc. Proc.,' vol. 48, p. 140 (1890).

† 'Zs. f. Sinnesphysiol.,' vol. 43, p. 76 (1908–9).

‡ *Ibid.*, vol. 42, pp. 24, 250 (1907); vol. 43, pp. 146, 199, 255 (1908–9).

§ 'Roy. Soc. Proc.,' B, vol. 86, p. 164 (1913).

In the focal plane of the telescope there was an adjustable slit, and behind this an extremely good nicol eyepiece, the position of which could be read on a divided circle. Both collimator slit and eyepiece slit were kept at a width of ½ a millimetre, and the telescope adjusted, so that the Na image formed by the light which missed the Wollaston prism fell exactly on the eyepiece slit. Superimposed on this image were the two formed by the beams which passed through the Wollaston prism. One of these was exactly the colour of Tl green (5351 A.U.); the other was a red of wave-length 6300 A.U. When the eyepiece slit was viewed through the eyepiece, it was seen illuminated by red, yellow, and green light superimposed. If, however, the eyepiece was removed, and the observer looked through the nicol and slit alone at the prism, he saw the face of the prism in Na yellow and superimposed on it the diamond-shaped Wollaston prism, the colour of the latter changing from red to green, through yellow, as the nicol was rotated. As viewed through the slit the face of the flint prism appeared almost square, each side subtending an angle of 6°; the diamond appeared in the middle of the square, and covered about one-third of its area.

The students tested comprised all those taking a course in the optical laboratory during the three terms of Session 1921–22, with one or two exceptions, owing to absence and some additions; 423 were men and 104 women. They were essentially a random distribution. They were taken in pairs from the experiment on which they were employed, optical bench, or whatever it was, and no one who was asked to make the test declined. Each made two settings, one in each of two adjacent quadrants; suppose, for example, these settings were 26·0° and 127·2°, which are very near the mean values, then the square of the cotangent of half the difference, $\cot^2 \frac{1}{2}(127{\cdot}2 - 26{\cdot}0) = 0{\cdot}6748$, gave the ratio of the intensity of red to intensity of green in the mixture, the intensities of red and green being measured in different units.

Certain precautions have to be observed in making the tests. First of all, at the wave-length selected for the yellow, Na yellow, the colour changes very rapidly with the wave-length, and a slight difference in the tint of the yellow means a large difference in the ratio of red to green. Owing to students seizing the telescope too roughly, these slight differences occurred occasionally, but they were always detected, as I read the instrument after every two students. Then the lamp used as source, a 40-watt Osram, must burn at the same colour throughout the tests. Before the investigation was started, three lamps of exactly the same colour were selected out of a batch of twenty; the first burned out and was replaced by the second half-way through the investigation. For the colour to be the same, the voltage must be constant; as the lamp was simply connected to the town leads the voltage was subject to small variations. I stopped the tests on two occasions, as it appeared the readings were changing appreciably owing to a slight change of voltage, but I could not be sure whether the change was a real one or not; however, it was thought better to err on the side of excessive caution.

If immediately before the test the eye is fatigued with red or green, a different setting is obtained. My usual value for the ratio of red to green is about 0·588. On fatiguing my eyes by gazing at a strong red light for 7 minutes and making a setting as rapidly as possible thereafter, I obtained 0·840, but the eye recovered fully in about 1 minute. On another occasion, on fatiguing my eyes in the same way, and reading as rapidly as possible, I obtained 0·729 instead of 0·588. This time my eyes took 2 minutes to recover. The room in which the tests were made was, however, darkened, and of no pronounced colour, so there was no occasion to suspect that the results were affected by fatigue.

Some of the 423 men were colour blind. These were asked to make a setting the same as the others. In every case they had a preference for one match, though in one or two cases they stated it was the best of a bad business, and one man required to be assisted by leading questions: it was, however, lack of self-confidence he was afflicted with rather than colour blindness. As it is very difficult to draw a line between the colour blind and the others, they have been entered in the Table and diagram along with the others.

Hitherto, it has been usual to state the results in terms of intensity of red light to intensity of green in the mixture, or *vice versa*. There is no particular reason why the results should be stated in terms of red to green rather than of green to red; both red and green enter into the mixture in precisely the same manner. The shape of the curve differs considerably according as red or green is in the numerator. I have thought it advisable, therefore, to state the results in terms of logarithm of ratio of intensity of red to green. When the logarithm of the ratio is taken as abscissa, inverting the ratio does not alter the shape of the diagram, but merely interchanges left and right.

The results are given in the following Table and in the diagram on p. 358:—

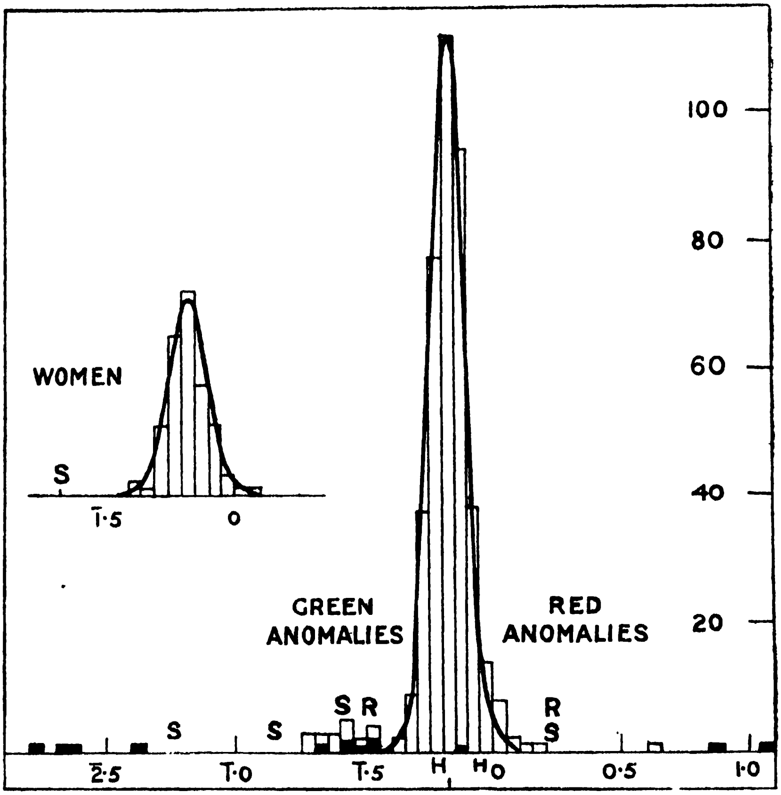
Logarithm of ratio of red to green.	Men.		Women.	
	Observed.	Calculated.	Observed.	Calculated.
$\bar{1}\cdot25$ and $<\bar{1}\cdot30$	3			
$\bar{1}\cdot30$ $<\bar{1}\cdot35$	3			
$\bar{1}\cdot35$ $<\bar{1}\cdot40$	3			
$\bar{1}\cdot40$ $<\bar{1}\cdot45$	5			
$\bar{1}\cdot45$ $<\bar{1}\cdot50$	2			
$\bar{1}\cdot50$ $<\bar{1}\cdot55$	4			
$\bar{1}\cdot55$ $<\bar{1}\cdot60$	0	0·2		
$\bar{1}\cdot60$ $<\bar{1}\cdot65$	2	1·9	2	0·5
$\bar{1}\cdot65$ $<\bar{1}\cdot70$	9	11·5	1	3
$\bar{1}\cdot70$ $<\bar{1}\cdot75$	37	41·8	11	11
$\bar{1}\cdot75$ $<\bar{1}\cdot80$	77	87	25	23
$\bar{1}\cdot80$ $<\bar{1}\cdot85$	112	112	32	29·5
$\bar{1}\cdot85$ $<\bar{1}\cdot90$	94	87	17	23
$\bar{1}\cdot90$ $<\bar{1}\cdot95$	38	41·8	11	11
$\bar{1}\cdot95$ $<0\cdot00$	14	11·5	3	3
$0\cdot00$ $<0\cdot05$	8	1·9	1	0·5
$0\cdot05$ $<0\cdot10$	2	0·2	1	
$0\cdot10$ $<0\cdot15$	1			
$0\cdot15$ $<0\cdot20$	1			

There were also men at $\bar{3}\cdot079$, $\bar{2}\cdot230$, $\bar{2}\cdot301$, $\bar{2}\cdot370$, $\bar{2}\cdot633$, 0·624, 0·876, 1·085.

We see from the Table, for example, that 77 men obtained a ratio of red to green, the logarithm of which lay in the range $\bar{1}\cdot75$ to $\bar{1}\cdot80$, including the lower limit and excluding the upper. These are represented in the diagram by a rectangle bounded by vertical lines through $\bar{1}\cdot75$ and $\bar{1}\cdot80$, and 77 units high. In order to avoid extending the diagram too far the man at $\bar{3}\cdot079$ has not been represented.

A Gaussian curve was fitted to all the data for the women, as shown in the diagram. Its maximum is at $\bar{1}\cdot825$, and the calculated values shown in the Table are the mid-ordinates of each rectangle. The Gaussian curve shown in the diagram for the men, and the corresponding calculated values are obtained from the curve for the women by increasing all the ordinates in a given ratio; a better agreement might have been obtained by fitting a curve directly to the data for the men, but there would have been a certain amount of arbitrariness in deciding how many of the outlying observers to include.

The question arises as to how much of the "scatter" of the Gaussian curve is due to error of observation, inexperience in making the readings, accidental variations, etc., and how much is due to a real difference in the physiological equipment of the observer. Each observer made two settings which meant two independent determinations; by taking the square of the difference and summing for a large number of observers, it was possible to calculate the



" scatter " which would be obtained if all the readings were made by an average observer, and it seems that about one-half of the breadth of the Gaussian curve is due to these accidental variations. This is a much larger proportion than I expected before making the calculation. My own setting is indicated by H_1; by fatiguing my eyes as much as possible with red light I got the setting H. Never under ordinary circumstances do I get the reading out by more than half the breadth of a rectangle either way. But some observers are much more erratic than others.

None of the women investigated showed any abnormality of colour vision. Of the men the following ten were very colour blind: (92) 2·663, (137) $\bar{2}$·301, (144) $\bar{3}$·079, (145) 0·876, (149) $\bar{2}$·230, (233) 1·085, (269) $\bar{1}$·516, (350) $\bar{1}$·894, (357) $\bar{1}$·406, (358) $\bar{2}$·370; the number in the bracket is the observer's number, the other is the logarithm of his ratio. The following four were partially colour blind: (244) $\bar{1}$·349, (250) $\bar{1}$·547, (274) $\bar{1}$·433, (405) $\bar{1}$·452. The foregoing are entered in black in the diagram. The following were put through a very stringent test, as it was thought, owing to their setting, that their colour vision might be defective, but the test showed they were as good as the normal: (105) 1·255, (293) $\bar{1}$·403, (339) 0·624, (342) $\bar{1}$·361, (411) $\bar{1}$·338, (104) $\bar{1}$·336.

We come now to consider the results. There is no doubt whatever of the existence of small differences following the ordinary law of error, and it is clear that the terms green anomaly and red anomaly have hitherto been used in two senses. Edridge-Green found 12 anomalies among 100 women students. My 104 women students all group themselves according to the ordinary law of error. Edridge-Green consequently applies the term to the outlying portions of the normal curve. Schuster, on the contrary, restricts the term anomalous trichromat to the scattered observers outside the normal curve. Schuster had, altogether, 3 colour blind and 5 anomalous trichromats out of 75 observers. I have 14 colour blind and 15 others outside the normal curve out of 423 men students. The proportion agrees as well as can be expected. I have indicated the positions of Schuster's anomalous trichromats by S, and two of Rayleigh's anomalous trichromats by R in the diagrams; these positions are, however, only approximate, as Rayleigh and Schuster used a different red and green from mine.

With reference to the women, if it were not for Schuster's observer, I should have said that normal variation covered everything. There are about 28 men outside the normal curve; if the proportion were the same, there should be about 7 women; but there are none. It is clear, therefore, that normal variation almost covers everything as regards the women. As regards the men, I have not yet found a colour blind man who is a dichromat, though I have now worked through a population of 1500; trichromasy, in my experience, seems to approach monochromasy directly without passing through dichromasy as an intermediate stage.* I consequently do not like the term anomalous trichromats as a designation for those observers outside the normal curve who are not colour blind, but prefer the terms red and green anomalies. Using the terms in this sense, I have thus one or two red anomalies and 14 green anomalies. Although some of these have remarkably good colour vision, it is not easy to draw a sharp distinction between them and the colour blind; there are intermediate cases.

There is no doubt that the anomalies cannot be explained as extreme cases of normal variation. But they are not nearly so common as would be anticipated from Rayleigh's first paper. And they do not appear to form well defined classes.

Some observers see the red end of the spectrum much dimmer than others, and recently a survey has been started to see whether these are extreme cases of normal variation, or an independent group. After preliminary work with a spectrometer, wedge photometer, and flicker photometer, it was found that matching a red lamp against a white light by means of a photometer of the grease spot type gave the most satisfactory results. The red light was spectrally very pure. This survey has overlapped the one described in this paper to the extent of about 120 observers, and the following Table includes all the colour blind and red and green anomalies of the present paper who have also made the photometer test. The first column gives the observer's number, the second the logarithm of his ratio of red to green, and the third the photometer reading; a low reading means that the observer is very blind to red. The scale has not yet been calibrated, but the observer (357), at the one end, sees red about 60 times as bright as the observer (149), at the other. The normal setting on the photometer scale is about 18; the Table includes the two most red blind and two most red sensitive observers yet disclosed by the photometer survey :—

---

* 'Roy. Soc. Edin. Proc.,' vol. 42, p. 75 (1922).

(149)	$\bar{2}{\cdot}230$	<0	colour blind.
(293)	$\bar{1}{\cdot}403$	16·5	green anomaly.
(334)	1·376	15·6	green anomaly.
(339)	0·624	1·68	red anomaly.
(342)	$\bar{1}{\cdot}361$	12·5	green anomaly.
(350)	$\bar{1}{\cdot}894$	16·0	colour blind.
(357)	$\bar{1}{\cdot}406$	27·4	colour blind.
(358)	$\bar{2}{\cdot}370$	25·0	colour blind.
(405)	$\bar{1}{\cdot}452$	19·4	partially colour blind.
(411)	$\bar{1}{\cdot}338$	19·2	green anomaly.

If we examine the Table, we find one very colour blind man (350), normal in both tests; one colour blind (149), and one red anomaly (339), extremely red blind. Four green anomalies (293), (334), (342), (411), and one partially colour blind man (405), who makes the same reading as these four in the Rayleigh test, about normal in the photometer test; and two colour blind (357) and (358), extremely red sensitive. So the probability is that those who are abnormal in the Rayleigh test will be abnormal also in the photometer test, but there are marked exceptions; also, from the nature of the abnormality in the one test, it is not possible to predict the nature of the abnormality in the other.

The Line of Action of the Resultant Pressure in Discontinuous Fluid Motion.

By S. BRODETSKY, M.A., Ph.D., University of Leeds.

(Communicated by Sir George Greenhill, F.R.S. Received June 7, 1922.)

1. The problem of any barrier in a fluid stream is best attacked by the method due to Levi-Civita,* of which useful accounts, with extensions, are given by Cisotti† and Brillouin.‡ The resultant pressure for any barrier has been given in terms of the constants defining the barrier; but the calculations required to find the line of action of this pressure have not been carried out. It is our object to supply this deficiency here.

The motion is two-dimensional. Let the complex variable $z (\equiv x + \iota y)$ define position in any plane perpendicular to the generators of the barrier, the x axis being parallel to the direction of the stream at infinity. We define

$$u = \frac{\partial \phi}{\partial x} = \frac{\partial \psi}{\partial y}, \quad v = \frac{\partial \phi}{\partial y} = -\frac{\partial \psi}{\partial x},$$

where u, v are the velocity components, and ϕ, ψ are the velocity potential and stream function respectively. Let $w \equiv \phi + \iota\psi$ and define ζ, Ω, r, θ so that

$$\zeta \equiv re^{\iota\theta} = \frac{dz}{dw}; \; \Omega = \log \zeta \equiv \log r + \iota\theta. \tag{1}$$

The figure shows the z, w, ζ, Ω planes for such a problem, with the boundary conditions. C is the point of bifurcation of the stream-line IC, CA, CA′ are the stream-lines in contact with the barrier, AJ, A′J′ are the free stream-lines. We take the standard dimensions to be unit velocity at infinity unit value of $\frac{1}{2}(\sqrt{\text{CA}} + \sqrt{\text{CA}'})$ as measured in the w plane, and unit density. For other values of these quantities we multiply lengths in the z plane by the square of $\frac{1}{2}(\sqrt{\text{CA}} + \sqrt{\text{CA}'})$ in the w plane, and we multiply the resultant pressure per unit length of the barrier by the square of the velocity at infinity, by the density, and by the square of $\frac{1}{2}(\sqrt{\text{CA}} + \sqrt{\text{CA}'})$ in the w plane.

We introduce the transformation in which the variable $\tau \equiv \rho e^{\iota\sigma}$ is given by

$$\sqrt{w} = \frac{1}{2\iota}\left(\tau - \frac{1}{\tau}\right) - \sin \sigma_0, \tag{2}$$

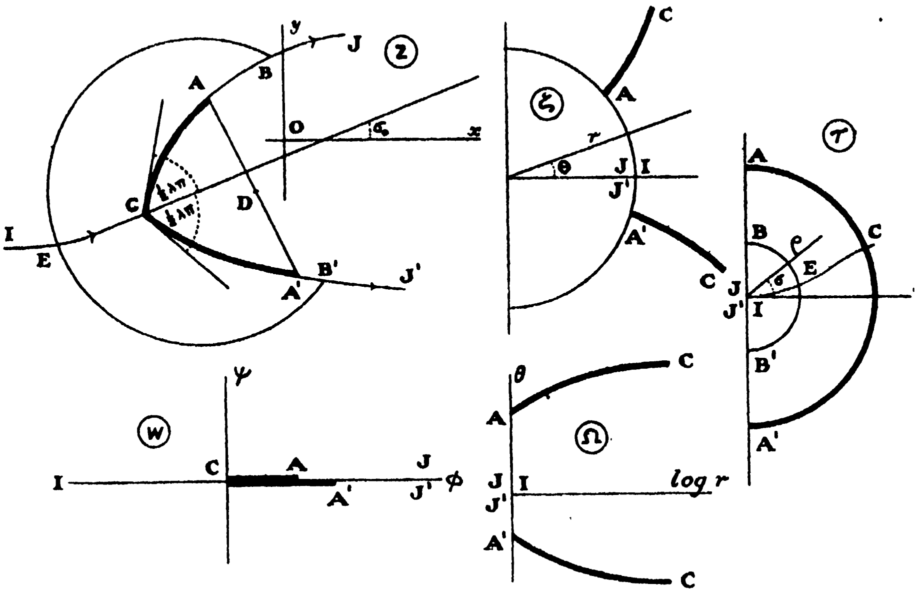
where the point C is given by $\tau = e^{\iota\sigma_0}$. It is immediately verified that the

---

* 'Rend. del Circ. di Palermo,' vol. 23 (November, 1906).

† 'Idromeccanica Piana,' vol. 2, Milano (1922).

‡ 'Ann. de Chim. et Phys.,' vol. 23, pp. 145–230 (1911).



barrier ACA′ is the semicircle $\rho = 1$, $-\pi/2 \leq \sigma \leq +\pi/2$ in the τ plane, while the free stream-lines AJ, A′J′ become the two halves of the diameter along the imaginary axis, namely $\tau = \iota\rho$, $-1 \leq \rho \leq +1$.

If the barrier is plane we have

$$\Omega = \log\left(\frac{1+e^{\iota\sigma_0\tau}}{1-e^{-\iota\sigma_0\tau}}\right),\ \zeta = \frac{1+e^{\iota\sigma_0\tau}}{1-e^{-\iota\sigma_0\tau}}. \qquad (3)$$

If it consists of two planes meeting at an angle $\lambda\pi$ at C (measured away from the streaming fluid), then

$$\Omega = \lambda\log\left(\frac{1+e^{\iota\sigma_0\tau}}{1-e^{-\iota\sigma_0\tau}}\right),\quad \zeta = \left(\frac{1+e^{\iota\sigma_0\tau}}{1-e^{-\iota\sigma_0\tau}}\right)^{\lambda}. \qquad (4)$$

In the general problem we write

$$\Omega = \lambda\log\left(\frac{1+e^{\iota\sigma_0\tau}}{1-e^{-\iota\sigma_0\tau}}\right) + A_1\tau + \tfrac{1}{2}\iota A_2\tau^2 + \tfrac{1}{3}A_3\tau^3 + \tfrac{1}{4}\iota A_4\tau^4 + \ldots, \qquad (5)$$

$$\equiv a_1\tau + \tfrac{1}{2}\iota a_2\tau^2 + \tfrac{1}{3}a_3\tau^3 + \tfrac{1}{4}\iota a_4\tau^4 + \ldots, \qquad (6)$$

where the A's, a's are all real, and the expansion in (5) is convergent for $|\tau| \leq 1$. The angle between the tangents at C in the z plane is again $\lambda\pi$, while the bisector of this angle makes with the x axis the angle (marked σ_0):—

$$\lambda\sigma_0 + A_1\sin\sigma_0 + \tfrac{1}{2}A_2\cos 2\sigma_0 + \tfrac{1}{3}A_3\sin 3\sigma_0 + \tfrac{1}{4}A_4\cos 4\sigma_0 + \ldots.$$

If then we have any set of constants we can find the barrier and its position.

2. In our notation the components of pressure on unit length of the barrier, R_x, R_y, and the moment M about the origin in the z plane, are given by

$$\left.\begin{aligned} R_y+\iota R_x &= \tfrac{1}{2}\int\limits_{B'EB}\frac{dz}{\zeta^2}-\tfrac{1}{2}\int\limits_{B'EB}(dx-\iota dy), \\ M &= \textit{real part of}\left\{\tfrac{1}{2}\int\limits_{B'EB}\frac{z\,dz}{\zeta^3}-\tfrac{1}{4}\Big[\textstyle\int(dx+\iota dy)\int(dx-\iota dy)\Big]_{B'}^{B}\right\} \end{aligned}\right\} \qquad (7)$$

the integration referring to any contour like B′EB, where B′ is on A′J′, B is on AJ, and the contour lies wholly within the moving fluid. If we express the integrals in terms of τ we can take for our contour any curve like B′EB (see τ plane in the figure). It is well known that

$$R_x=\frac{\pi}{4}a_1^2,\ R_y=\frac{\pi}{4}(a_2+4a_1\sin\sigma_0), \qquad (8)$$

(see Cisotti, p. 175; Brillouin, p. 195).

As regards the moment Cisotti does not go beyond giving a formula like (7), while Brillouin makes a few general remarks: neither works out the explicit value of M, and neither defines the meaning of the term "origin" in this connection.

Since the solution (5, 6) gives us really the value of $dz/d\tau$, and we have to integrate in order to get z in terms of τ, it is clear that no origin has yet been defined in the z plane. We shall define the origin as follows. Let $dz/d\tau$ be expanded in terms of τ; we shall soon see that the expansion contains terms in $\tau^{-3}, \tau^{-2}, \tau^{-1}, \tau^0, \tau^1, \tau^2$ .... Integrate the series without introducing any arbitrary constant, so that we get terms in $\tau^{-2}, \tau^{-1}, \log\tau, \tau^1, \tau^2, \tau^3$ .... We take the origin to correspond to the value of τ that makes the expansion thus obtained vanish.

3. If we use the fact that $dz/dw=\zeta$, and let the symbol $(\quad)_-$ denote that in the expression enclosed by the brackets, $-\iota$ is to be substituted for $+\iota$, we can write M in the form

M = *real part of*

$$\left\{\tfrac{1}{2}\int\limits_{B'EB}\frac{1}{\zeta}\frac{dw}{d\tau}\left(\int\zeta\frac{dw}{d\tau}\,d\tau\right)d\tau-\tfrac{1}{4}\left[\left(\int\zeta\frac{dw}{d\tau}\,d\tau\right)\left(\int\zeta\frac{dw}{d\tau}\,d\tau\right)_-\right]_{B'}^{B}\right\} \qquad (9)$$

But

$$\zeta=\exp.(a_1\tau+\tfrac{1}{2}\iota a_2\tau^2+\tfrac{1}{3}a_3\tau^3+\ \ldots),$$

so that $1/\zeta$ is obtained from ζ by changing the signs of all the a's. Also

$$\frac{dw}{d\tau}=\frac{1+\tau^2}{2\tau^3}(1+2\iota\sin\sigma_0.\tau-\tau^2)=\frac{1+\tau^2}{2\tau^3}(1+e^{\iota\sigma_0}\tau)(1-e^{-\iota\sigma_0}\tau). \qquad (10)$$

Hence it follows that $\zeta dw/d\tau$ begins with a term τ^{-3}. Further, if we use the notation

$$\zeta \frac{dw}{d\tau} = -\frac{2(\alpha_{-2}+\iota\beta_{-2})}{\tau^3} - \frac{(\alpha_{-1}+\iota\beta_{-1})}{\tau^2} + \frac{(\alpha_0+\iota\beta_0)}{\tau} + (\alpha_1+\iota\beta_1) + 2(\alpha_2+\iota\beta_2)\tau + \ldots, \tag{11}$$

we at once see that $\zeta^{-1}dw/d\tau$ is got by changing the signs of β, α alternately, *i.e.*,

$$\frac{1}{\zeta}\frac{dw}{d\tau} = -\frac{2(\alpha_{-2}-\iota\beta_{-2})}{\tau^3} + \frac{(\alpha_1-\iota\beta_{-1})}{\tau^2} + \frac{(\alpha_0-\iota\beta_0)}{\tau} - (\alpha_1-\iota\beta_1) + 2(\alpha_2-\iota\beta_2)\tau - \ldots. \tag{12}$$

Again, by the definition of the origin, (11) gives

$$\int \zeta \frac{dw}{d\tau}\,d\tau = \frac{(\alpha_{-2}+\iota\beta_{-2})}{\tau^2} + \frac{(\alpha_{-1}+\iota\beta_{-1})}{\tau} + (\alpha_0+\iota\beta_0)\log\tau + (\alpha_1+\iota\beta_1)\tau + (\alpha_2+\iota\beta_2)\tau^2 + \ldots \tag{13}$$

$$= \frac{(\alpha_{-2}+\iota\beta_{-2})}{\rho^2}e^{-2\iota\sigma} + \frac{(\alpha_{-1}+\iota\beta_{-1})}{\rho}e^{-\iota\sigma} + (\alpha_0+\iota\beta_0)(\log\rho+\iota\sigma) + (\alpha_1+\iota\beta_1)\rho e^{\iota\sigma} + (\alpha_2+\iota\beta_2)\rho^2 e^{2\iota\sigma} + \ldots \tag{14}$$

Hence we deduce that

$$\left(\int \zeta \frac{dw}{d\tau}\,d\tau\right)_- = \frac{(\alpha_{-2}-\iota\beta_{-2})}{\rho^2}e^{2\iota\sigma} + \frac{(\alpha_1-\iota\beta_{-1})}{\rho}e^{\iota\sigma} + (\alpha_0-\iota\beta_0)(\log\rho-\iota\sigma) + (\alpha_1-\iota\beta_1)\rho e^{-\iota\sigma} + (\alpha_2-\iota\beta_2)\rho^2 e^{-2\iota\sigma} + \ldots \tag{15}$$

If we form the product of (14), (15) we get terms in ρ^{-4}, ρ^{-3}, ρ^{-2}, ρ^{-1}, $\log\rho$, ρ^0, ρ^1, ρ^2, ...

Now let us take the contour B′EB in the τ plane to be a semicircle whose radius ρ is very small, so that $\rho \to 0$. We can then omit all terms with positive powers of ρ. Also the negative powers of ρ and the $\log\rho$ term must necessarily cut out with the corresponding terms in the other part of M. Thus in the product $\left(\int \zeta \frac{dw}{d\tau}\,d\tau\right)\left(\int \zeta \frac{dw}{d\tau}\,d\tau\right)_-$ we need only take the terms independent of ρ, *i.e.*,

$$2(\alpha_2\alpha_{-2}+\beta_2\beta_{-2})\cos 4\sigma + 2(\alpha_2\beta_{-2}-\alpha_{-2}\beta_2)\sin 4\sigma + 2(\alpha_1\alpha_{-1}+\beta_1\beta_{-1})\cos 2\sigma + 2(\alpha_1\beta_{-1}-\alpha_{-1}\beta_1)\sin 2\sigma + (\alpha_0^2+\beta_0^2)\sigma^2;$$

and if we put successively $\sigma_0 = \pi/2$, $\sigma_0 = -\pi/2$ and subtract, we get zero.

To get the other part of M we have to multiply (12) into (13), and then integrate. We get terms in τ^{-5}, τ^{-4}, τ^{-3}, τ^{-2}, τ^{-1}, τ^0, τ^1, τ^2, ...; $\tau^{-3}\log\tau$, $\tau^{-2}\log\tau$, $\tau^{-1}\log\tau$, $\log\tau$, $\tau\log\tau$, $\tau^2\log\tau$, ... When we integrate we get terms in τ^{-4}, τ^{-3}, τ^{-2}, τ^{-1}, $\log\tau$, τ^1, τ^2, ...; $\tau^{-2}(\log\tau+\frac{1}{2})$, $\tau^{-1}(\log\tau+1)$, $(\log\tau)^2$, $\tau(\log\tau-1)$, ... If now we make $\tau=\rho e^{\iota\sigma}$ and let ρ be constant and small, so that $\rho\rightarrow 0$, the positive powers of ρ can be omitted, and the negative powers of ρ as well as the $\log\rho$ parts of the terms $\log\tau$, $(\log\tau)^2$ must be cancelled by corresponding terms in the part of M already considered. This can in fact be verified straightforwardly, but it is not worth while reproducing the rather complicated algebra here. The terms we have to consider are thus seen to be

$$[4\iota(\alpha_2\beta_{-2}-\alpha_{-2}\beta_2)-2\iota(\alpha_1\beta_{-1}-\alpha_{-1}\beta_1)]\log\tau+(\alpha_0^2+\beta_0^2)(\log\tau)^2,$$

and we find that

$$\begin{aligned}\mathrm{M}&=\left[\{(\alpha_1\beta_{-1}-\alpha_{-1}\beta_1)-2(\alpha_2\beta_{-2}-\alpha_{-2}\beta_2)\}\,\sigma-(\alpha_0^2+\beta_0^2)\,\sigma^2\right]_{-\pi/2}^{\pi/2}\\&=\pi[(\alpha_1\beta_{-1}-\alpha_{-1}\beta_1)-2(\alpha_2\beta_{-2}-\alpha_{-2}\beta_2)].\end{aligned}$$

The form of M shows that we need find only the first five terms in (11), so that we have to expand ζ only as far as τ^4. If we do this and substitute in M we get

$$\mathrm{M}=\frac{\pi}{32}\{3a_1^2a_2-a_4+4\sin\sigma_0(a_1^3+2a_1-2a_3)+16a_2\sin^2\sigma_0\}\ \ldots \tag{16}$$

where, in terms of the A's we have

$$\begin{aligned}a_1&=\mathrm{A}_1+2\lambda\cos\sigma_0,\ a_2=\mathrm{A}_2-2\lambda\sin 2\sigma_0,\\a_3&=\mathrm{A}_3+2\lambda\cos 3\sigma_0,\ a_4=\mathrm{A}_4-2\lambda\sin 4\sigma_0,\ \ldots.\end{aligned}$$

4. If the barrier is plane then

$$\lambda=1,\ \mathrm{A}_1=\mathrm{A}_2=\ldots=0;$$

so that by (8), (16) we have

$$\mathrm{R}_x=\pi\cos^2\sigma_0,\ \mathrm{R}_y=\pi\sin\sigma_0\cos\sigma_0,\ \mathrm{M}=-\frac{\pi}{4}\sin\sigma_0\cos\sigma_0.$$

The resultant pressure is therefore $\pi\cos\sigma_0$ perpendicular to the plane, and its distance from the origin is

$$\tfrac{1}{4}\sin\sigma_0$$

in such a sense that the moment is negative, *i.e.*, above the origin in our figure when σ_0 is positive. We can find the origin quite easily. By (3), (10)

$$\frac{dz}{d\tau}=\frac{1+\tau^2}{2\tau^3}(1+e^{\iota\sigma_0}\tau)^2,$$

and the origin is given by that value of τ for which

$$\tfrac{1}{4}\left(e^{\iota\sigma_0}\tau^2-\frac{e^{-\iota\sigma_0}}{\tau^2}\right)+\left(\tau-\frac{1}{\tau}\right)+\cos\sigma_0\,.\log\tau=0,$$

so that $\rho=1$ and σ is given by the equation

$$\tfrac{1}{2}\sin(2\sigma+\sigma_0)+2\sin\sigma+\sigma\cos\sigma_0=0.$$

The origin is thus on the barrier when it is plane.

It is not convenient to solve the equation for σ, and in fact this is not necessary. The general value of z in terms of τ is

$$ze^{-\iota\sigma_0} = \tfrac{1}{4}\left(e^{\iota\sigma_0}\tau^2 - \frac{e^{-\iota\sigma_0}}{\tau^2}\right) + \left(\tau - \frac{1}{\tau}\right) + \cos\sigma_0 . \log\tau.$$

Hence for the points A, A′, at which $\tau = +\iota$, $\tau = -\iota$ respectively, we have

$$z_A = \iota e^{\iota\sigma_0}\left(2 - \tfrac{1}{2}\sin\sigma_0 + \frac{\pi}{2}\cos\sigma_0\right), \quad z_{A'} = -\iota e^{\iota\sigma_0}\left(2 + \tfrac{1}{2}\sin\sigma_0 + \frac{\pi}{2}\cos\sigma_0\right).$$

It follows that the width of the barrier is

$$4 + \pi\cos\sigma_0,$$

while the distance of the origin from the mid-line of the barrier is

$$\tfrac{1}{2}\sin\sigma_0$$

towards A. Hence the resultant pressure meets the barrier at distance

$$\tfrac{3}{4}\sin\sigma_0$$

from the mid-line measured towards A. If then the width of the plane barrier is l, the distance of the centre of pressure from the mid-line is

$$\tfrac{3}{4}l\,\frac{\sin\sigma_0}{4 + \pi\cos\sigma_0},$$

the well-known result.

It is of interest to note that the ratio of the distances of the origin and of the centre of pressure from the mid-line is 2 : 3 for all values of σ_0. When $\sigma_0 = 0$, the plane barrier being normal to the stream, the two points coincide on the mid-line.

5. In more general cases the expansion of $\zeta dw/d\tau$ in terms of τ is often impracticable, and we proceed as follows. Using the notation of (11) we see that

$$\int\left\{\zeta\frac{dw}{d\tau} + \frac{2(\alpha_{-2}+\iota\beta_{-2})}{\tau^3} + \frac{(\alpha_{-1}+\iota\beta_{-1})}{\tau^2} - \frac{(\alpha_0+\iota\beta_0)}{\tau}\right\}d\tau$$

vanishes when $\tau = 0$. Hence its value for any point τ can be written

$$\int_0^\tau\left\{\zeta\frac{dw}{d\tau} + \frac{2(\alpha_{-2}+\iota\beta_{-2})}{\tau^3} + \frac{(\alpha_{-1}+\iota\beta_{-1})}{\tau^2} - \frac{(\alpha_0+\iota\beta_0)}{\tau}\right\}d\tau.$$

Thus the value of z in terms of τ without any arbitrary constants is

$$z = \frac{(\alpha_{-2}+\iota\beta_{-2})}{\tau^2} + \frac{(\alpha_{-1}+\iota\beta_{-1})}{\tau} + (\alpha_0+\iota\beta_0)\log\tau$$
$$+\int_0^\tau\left\{\zeta\frac{dw}{d\tau} + \frac{2(\alpha_{-2}+\iota\beta_{-2})}{\tau^3} + \frac{(\alpha_{-1}+\iota\beta_{-1})}{\tau^2} - \frac{(\alpha_0+\iota\beta_0)}{\tau}\right\}d\tau \;\ldots\ldots \quad (17)$$

Now the integrand in the integral in (17) has no singularities inside the circle $|\tau| = 1$, and on the circumference of this circle the only possible singularity is at C, *i.e.*, $\tau = e^{\iota\sigma_0}$, when $\lambda > 1$. Hence if τ is any point inside or on the circumference of this circle, except C, we can choose for the path of integration in (17) any contour lying within the semicircle defined in the τ plane, and joining the origin $\tau = 0$ to the point τ. Let us choose the straight line lying between the origin $\tau = 0$ and the point τ. If τ is a point on the circumference of $|\tau| = 1$, so that $\rho = 1$, $\tau = e^{\iota\sigma}$, we get

$$z = (\alpha_{-2}+\iota\beta_{-2})\,e^{-2\iota\sigma}+(\alpha_{-1}+\iota\beta_{-1})\,e^{-\iota\sigma}+\iota\,(\alpha_0+\iota\beta_0)\,\sigma$$
$$+\int_0^1\left\{\left(\zeta\frac{dw}{d\tau}\right)_{\tau=\rho e^{\iota\sigma}}+\frac{2\,(\alpha_{-2}+\iota\beta_{-2})}{\rho^3}\,e^{-3\iota\sigma}+\frac{(\alpha_{-1}+\beta_{-1})}{\rho_2}\,e^{-2\iota\sigma}-\frac{(\alpha_0+\iota\beta_0)}{\rho}\,e^{-\iota\sigma}\right\}e^{\iota\sigma}.d\rho,$$

where z is now a point on the barrier ACA′ in the z plane.

Let $\sigma = \pm\pi/2$, corresponding to the points A, A′ respectively, and use the value of $dw/d\tau$ given by (10): we find

$$z_A = -(\alpha_{-2}+\iota\beta_{-2})-\iota\,(\alpha_{-1}+\iota\beta_{-1})+\frac{\iota}{2}\,\pi\,(\alpha_0+\iota\beta_0)$$
$$+\int_0^1\left\{-\frac{1-\rho^2}{2\rho^3}(1-2\rho\sin\sigma_0+\rho^2)\,\zeta(\iota\rho)\right.$$
$$\left.-\frac{2\,(\alpha_{-2}+\iota\beta_{-2})}{\rho^3}-\frac{\iota\,(\alpha_{-1}+\iota\beta_{-1})}{\rho^2}-\frac{(\alpha_0+\iota\beta_0)}{\rho}\right\}d\rho,$$

$$z_{A'} = -(\alpha_{-2}+\iota\beta_{-2})+\iota\,(\alpha_{-1}+\iota\beta_{-1})-\frac{\iota}{2}\,\pi\,(\alpha_0+\iota\beta_0)$$
$$+\int_0^1\left\{-\frac{1-\rho^2}{2\rho^3}\,(1+2\rho\sin\sigma_0+\rho^2)\,\zeta(-\iota\rho)\right.$$
$$\left.-\frac{2\,(\alpha_{-2}+\iota\beta_{-2})}{\rho^3}+\frac{\iota(\alpha_{-1}+\iota\beta_{-1})}{\rho^2}-\frac{(\alpha_0+\iota\beta_0)}{\rho}\right\}d\rho,$$

where $\zeta(\iota\rho)$, $\zeta(-\iota\rho)$ are the values of ζ for $\tau = \pm\iota\rho$ respectively. The straight line joining AA′ in the z plane can be appropriately called the *chord*. If D is the mid-point of AA′, defining the *mid-line* of the barrier, we have at once

$$z_D = -(\alpha_{-2}+\iota\beta_{-2})$$
$$+\int_0^1\left\{-\frac{1-\rho^2}{4\rho^3}\,[(1-2\rho\sin\sigma_0+\rho^2)\,\zeta(\iota\rho)+(1+2\rho\sin\sigma_0+\rho^2)\,\zeta(-\iota\rho)]\right.$$
$$\left.-\frac{2\,(\alpha_{-2}+\iota\beta_{-2})}{\rho^3}-\frac{(\alpha_0+\iota\beta_0)}{\rho}\right\}d\rho.$$

A little algebra shows that

$$z_D = \tfrac{1}{4}-\int_0^1\left\{\frac{1-\rho^2}{4\rho^3}\,[(1-2\rho\sin\sigma_0+\rho^2)\,\zeta(\iota\rho)+(1+2\rho\sin\sigma_0+\rho^2)\,\zeta(-\iota\rho)]\right.$$
$$\left.-\frac{1}{2\rho^3}+\frac{a_1^2+\iota\,(a_2+4a_1\sin\sigma_0)}{4\rho}\right\}d\rho,\ldots \qquad (18)$$

and that the value of the integrand in the integral of (18) is zero when $\rho=0$.

If the function $\zeta(\tau)$ can be conveniently divided into an even function of τ and an odd function of τ, let these be $Z_0(\tau)$, $Z_1(\tau)$ respectively. Then

$$Z_0(-\iota\rho) = Z_0(\iota\rho), \qquad Z_1(-\iota\rho) = -Z_1(\iota\rho),$$

and equation (18) takes the form

$$\begin{aligned} z_D = \tfrac{1}{4} - \int_0^1 \Bigg\{ & \frac{1-\rho^4}{2\rho^3} Z_0(\iota\rho) \\ & - \frac{1-\rho^2}{\rho^2} \sin\sigma_0 \,.\, Z_1(\iota\rho) - \frac{1}{2\rho^3} + \frac{a_1^2 + \iota(a_2 + 4a_1 \sin\sigma_0)}{4\rho} \Bigg\} d\rho. \end{aligned} \qquad (19)$$

The co-ordinates x_D, y_D of the point D are given by (18) or (19), whichever is more convenient. Thus for the plane barrier we use (3), and we get

$$z_D = \tfrac{1}{4}(1 - e^{2\iota\sigma_0}) = \tfrac{1}{2} \sin\sigma_0 (\sin\sigma_0 - \iota \cos\sigma_0),$$

agreeing with the results already obtained.

In more complicated problems we can find x_D, y_D by means of numerical or graphical integration, remembering that the integrand in the integral of (18) or (19) is zero when $\rho = 0$.

6. Again, by the equations of § 5 we have

$$\begin{aligned} \tfrac{1}{2}(z_A - z_{A'}) = & -(\alpha_1 + \iota\beta_1) + \frac{\iota}{2}\pi(\alpha_0 + \iota\beta_0) \\ & + \int_0^1 \Bigg\{ -\frac{1-\rho^2}{4\rho^3} [(1 - 2\rho \sin\sigma_0 + \rho^2)\zeta(\iota\rho) - (1 + 2\rho\sin\sigma_0 + \rho^2)\zeta(-\iota\rho)] \\ & - \frac{\iota(\alpha_{-1} + \iota\beta_{-1})}{\rho^2} \Bigg\} d\rho. \end{aligned}$$

Hence we find

$$\begin{aligned} \tfrac{1}{2}(z_A - z_{A'}) &= z_A - z_D = z_D - z_{A'} \\ &= -(\sin\sigma_0 - \frac{\iota}{2} a_1) - \frac{\pi}{8}(a_2 + 4a_1 \sin\sigma_0 - \iota a_1^2) \\ & \quad - \int_0^1 \Bigg\{ \frac{1-\rho^2}{4\rho^3} [(1 - 2\rho\sin\sigma_0 + \rho^2)\zeta(\iota\rho) - (1 + 2\rho\sin\sigma_0 + \rho^2)\zeta(-\iota\rho)] \\ & \quad + \frac{\sin\sigma_0 - \frac{\iota}{2} a_1}{\rho^2} \Bigg\} d\rho, \end{aligned} \qquad (20)$$

or

$$\begin{aligned} &= -(\sin\sigma_0 - \frac{\iota}{2} a_1) - \frac{\pi}{8}(a_2 + 4a_1 \sin\sigma_0 - \iota a_1^2) \\ & \quad - \int_0^1 \Bigg\{ \frac{1-\rho^4}{2\rho^3} Z_1(\iota\rho) - \frac{1-\rho^2}{\rho^2} \sin\sigma_0 \,.\, Z_0(\iota\rho) + \frac{\sin\sigma_0 - \frac{\iota}{2} a_1}{\rho^2} \Bigg\} d\rho \ldots \end{aligned} \qquad (21)$$

where Z_0, Z_1 have the same meanings as in § 5, and the value of the integrand in (20) or (21), when $\rho = 0$, is

$$\tfrac{1}{12}(6a_1^2 \sin\sigma_0 + 12 \sin\sigma_0 + 3a_1 a_2 + 6\iota a_2 \sin\sigma_0 - 2\iota a_3 - \iota a_1^3). \qquad (22)$$

We can thus find the positions of A, A′ relatively to D in the z plane, and also the origin. The barrier itself can be drawn by means of its curvature in terms of direction, and the problem is completely solved.

7. In some problems it may be of interest to have the distance of the line of action of the resultant pressure from the mid-line of the barrier. This is found as follows: Using R_x, R_y, M as already defined in § 2, the moment of the pressure about D is

$$M - x_D R_y + y_D R_x,$$

in the anti-clockwise sense in our figure. Hence, by (8), (16), the distance of the resultant pressure measured from D towards A is

$$\frac{\frac{1}{8}a_4 + (a_3 - a_1)\sin\sigma_0 - 2a_2\sin^2\sigma_0 - \frac{1}{2}a_1^3\sin\sigma_0 - \frac{3}{8}a_1^2 a_2 + (a_2 + 4a_1\sin\sigma_0)x_D - a_1^2 y_D.}{\{a_1^4 + (a_2 + 4a_1\sin\sigma_0)^2\}^{1/2}} \tag{23}$$

It may also be useful to know that the resultant pressure meets the line through D perpendicular to the direction of the stream at distance

$$\frac{1}{8}\frac{a_4}{a_1^2} + \left(\frac{a_3 - a_1}{a_1^2} - \frac{1}{2}a_1\right)\sin\sigma_0 - 2\frac{a_2}{a_1^2}\sin^2\sigma_0 - \frac{3}{8}a_2 + \frac{a_2 + 4a_1\sin\sigma_0}{a_1^2}x_D - y_D \tag{24}$$

from D, measured positive on the side of A.

8. In the case of the problem defined by (4), so that the barrier consists of two planes meeting at C at an angle of $\lambda\pi$, whose bisector makes an angle of σ_0 with the direction of the stream at infinity, we have

$$a_1 = 2\lambda\cos\sigma_0, \quad a_2 = -2\lambda\sin 2\sigma_0, \quad a_3 = 2\lambda\cos 3\sigma_0, \quad a_4 = -2\lambda\sin 4\sigma_0, \ldots.$$

Equations (8), (16) give

$$R_x = \pi\lambda^2\cos^2\sigma_0, \quad R_y = \pi\lambda\sin\sigma_0\cos\sigma_0,$$

$$M = \frac{1}{4}\pi\lambda\sin\sigma_0\cos\sigma_0\{2(1-\lambda^2)\cos^2\sigma_0 - 1\}.$$

Further, substituting for ζ in (18), (20), we find

$$x_D = \frac{1}{4} - \int_0^1 \left\{ \frac{1-\rho^2}{4\rho^3}\left[(1 - 2\rho\sin\sigma_0 + \rho^2)\cos\left(2\lambda\tan^{-1}\frac{\rho\cos\sigma_0}{1-\rho\sin\sigma_0}\right) \right.\right.$$
$$\left.+ (1 + 2\rho\sin\sigma_0 + \rho^2)\cos\left(2\lambda\tan^{-1}\frac{\rho\cos\sigma_0}{1+\rho\sin\sigma_0}\right)\right] - \frac{1}{2\rho^3}$$
$$\left. + \frac{\lambda^2\cos^2\sigma_0}{\rho}\right\} d\rho,$$

$$y_D = -\int_0^1 \left\{ \frac{1-\rho^2}{4\rho^3}\left[(1 - 2\rho\sin\sigma_0 + \rho^2)\sin\left(2\lambda\tan^{-1}\frac{\rho\cos\sigma_0}{1-\rho\sin\sigma_0}\right) \right.\right.$$
$$\left. - (1 + 2\rho\sin\sigma_0 + \rho^2)\sin\left(2\lambda\tan^{-1}\frac{\rho\cos\sigma_0}{1+\rho\sin\sigma_0}\right)\right]$$
$$\left. + \frac{\lambda\sin\sigma_0\cos\sigma_0}{\rho}\right\} d\rho,$$

$$\tfrac{1}{2}(x_{\mathrm{A}}-x_{\mathrm{A}'}) = -\sin\sigma_0-\tfrac{1}{2}\pi\lambda\sin\sigma_0\cos\sigma_0$$
$$-\int_0^1\left\{\frac{1-\rho^2}{4\rho^3}\left[(1-2\rho\sin\sigma_0+\rho^2)\cos\left(2\lambda\tan^{-1}\frac{\rho\cos\sigma_0}{1-\rho\sin\sigma_0}\right)\right.\right.$$
$$\left.\left.-(1+2\rho\sin\sigma_0+\rho^2)\cos\left(2\lambda\tan^{-1}\frac{\rho\cos\sigma_0}{1+\rho\sin\sigma_0}\right)\right]+\frac{\sin\sigma_0}{\rho_2}\right\}d\rho,$$

$$\tfrac{1}{2}(y_{\mathrm{A}}-y_{\mathrm{A}'}) = \lambda\cos\sigma_0+\tfrac{1}{2}\pi\lambda^2\cos^2\sigma_0$$
$$-\int_0^1\left\{\frac{1-\rho^2}{4\rho^3}\left[(1-2\rho\sin\sigma_0+\rho^2)\sin\left(2\lambda\tan^{-1}\frac{\rho\cos\sigma_0}{1-\rho\sin\sigma_0}\right)\right.\right.$$
$$\left.\left.+(1+2\rho\sin\sigma_0+\rho^2)\sin\left(2\lambda\tan^{-1}\frac{\rho\cos\sigma_0}{1+\rho\sin\sigma_0}\right)\right]-\frac{\lambda\cos\sigma_0}{\rho^2}\right\}d\rho,$$

the integrands in the integrals of $x_{\mathrm{D}}, y_{\mathrm{D}}$ being zero when $\rho = 0$, while those in $\tfrac{1}{2}(x_{\mathrm{A}}-x_{\mathrm{A}'})$, $\tfrac{1}{2}(y_{\mathrm{A}}-y_{\mathrm{A}'})$ have the values

$$\sin\sigma_0,\quad \tfrac{1}{8}(\lambda-\lambda^3)\cos 3\sigma_0-\tfrac{1}{2}(\lambda+\lambda^3)\cos\sigma_0$$

when $\rho=0$. The knowledge of these values of the integrands is useful when numerical or graphical integration has to be resorted to.

Starting with any point in the z plane as D, and with a given direction of the x axis, namely, that of the stream at infinity, x_{D}, y_{D} fix the origin; $\tfrac{1}{2}(x_{\mathrm{A}}-x_{\mathrm{A}'})$, $\tfrac{1}{2}(y_{\mathrm{A}}-y_{\mathrm{A}'})$ fix the points A, A′. The angles that CA, CA′ make with the positive direction of the x axis are

$$\sigma_0+\tfrac{1}{2}\lambda\pi,\quad \sigma_0-\tfrac{1}{2}\lambda\pi.$$

We can therefore draw CA, CA′, and the geometry of the z plane is completely determined. The line of action of the resultant pressure makes an angle

$$\tan^{-1}\left(\frac{\tan\sigma_0}{\lambda}\right)$$

with the x axis, and is at distance

$$\tfrac{1}{4}\sin\sigma_0\frac{2(1-\lambda^2)\cos^2\sigma_0-1}{(\sin^2\sigma_0+\lambda^2\cos^2\sigma_0)^{1/2}}$$

from the origin, in such a sense that the moment is anti-clockwise when this distance is positive.

9. A curved barrier of some interest is that defined by

$$\zeta=\left(\frac{1+e^{\iota\sigma_0\tau}}{1-e^{-\iota\sigma_0\tau}}\right)\left(\frac{1+t_1e^{\iota\sigma_1\tau}}{1-t_1e^{-\iota\sigma_1\tau}}\right)^{-\cos\sigma_0/\cos\sigma_1\sin\alpha_1},$$

where t_1, α_1, σ_0, σ_1 are connected by the relations

$$t_1=\tan\tfrac{1}{2}\alpha_1,\quad \sin\sigma\ =\sin\alpha_1\sin\sigma_1.$$

Since the value of λ is unity, the barrier is continuous at C. For different values of α_1, σ_1, we get first approximations to barriers in the form of elliptic cylinders of various eccentricities and in different orientations relatively to the direction of the stream at infinity. We now have

$$a_1 = 2\cos\sigma_0 - \frac{2\cos\sigma_0}{\cos\sigma_1 \sin\alpha_1} t_1 \cos\sigma_1,$$

$$a_2 = -2\sin 2\sigma_0 + \frac{2\cos\sigma_0}{\cos\sigma_1\sin\alpha_1} t_1^2 \sin 2\sigma_1, \ldots$$

and equations (8), (16) give

$$R_x = \pi\cos^2\sigma_0\left(\frac{\cos\alpha_1}{1+\cos\alpha_1}\right)^2, \quad R_y = \pi\sin\sigma_0\cos\sigma_0\left(\frac{\cos\alpha_1}{1+\cos\alpha_1}\right)^2,$$

$$M = -\tfrac{1}{4}\pi\sin\sigma_0\cos\sigma_0\frac{\cos^2\alpha_1}{(1+\cos\alpha_1)^4}(1+8\cos\alpha_1+\cos^2\alpha_1+4\sin^2\sigma_0);$$

while the line of action of the resultant pressure is at distance

$$-\tfrac{1}{4}\frac{\sin\sigma_0}{(1+\cos\alpha_1)^2}(1+8\cos\alpha_1+\cos^2\alpha_1+4\sin^2\sigma_0)$$

from the origin, in such a sense that the moment is anti-clockwise when this distance is positive.

The geometry of the z plane is given by the direction θ at any point of the barrier, and the corresponding radius of curvature R, both expressed in terms of σ as a parameter:—

$$\theta = \pm\frac{\pi}{2} + \sigma_0 - \frac{\cos\sigma_0}{\cos\sigma_1\sin\alpha_1}\left\{\tan^{-1}\left(\frac{t_1\sin(\sigma+\sigma_1)}{1+t_1\cos(\sigma+\sigma_1)}\right) + \tan^{-1}\left(\frac{t_1\sin(\sigma-\sigma_1)}{1-t_1\cos(\sigma-\sigma_1)}\right)\right\},$$

$\pm$ according as $\sigma \gtrless \sigma_0$, and

$$R = -\frac{2(1+\cos\overline{\sigma+\sigma_0})}{\cos\sigma_0\cos\alpha_1}(1+\sin\alpha_1\cos\overline{\sigma+\sigma_1})$$
$$(1-\sin\alpha_1\cos\overline{\sigma-\sigma_1})\left(\frac{1+\sin\alpha_1\cos\overline{\sigma+\sigma_1}}{1-\sin\alpha_1\cos\overline{\sigma-\sigma_1}}\right)^{-\cos\sigma_0/2\cos\sigma_1\sin\alpha_1}.$$

We also have

$$x_D = \tfrac{1}{4} - \int_0^1\left\{\frac{1-\rho^2}{4\rho^3}[(1-2\rho\sin\sigma_0+\rho^2)\cos\theta_1 + (1+2\rho\sin\sigma_0+\rho^2)\cos\theta_2] - \frac{1}{2\rho^3} + \frac{\cos^2\sigma_0}{\rho}\left(\frac{\cos\alpha_1}{1+\cos\alpha_1}\right)^2\right\}d\rho,$$

$$y_D = -\int_0^1\left\{\frac{1-\rho^2}{4\rho^3}[(1-2\rho\sin\sigma_0+\rho^2)\sin\theta_1 - (1+2\rho\sin\sigma_0+\rho^2)\sin\theta_2] + \frac{\sin\sigma_0\cos\sigma_0}{\rho}\left(\frac{\cos\alpha_1}{1+\cos\alpha_1}\right)^2\right\}d\rho,$$

where

$$\theta_1 = 2\tan^{-1}\frac{\rho\cos\sigma_0}{1-\rho\sin\sigma_0} - \frac{2\cos\sigma_0}{\cos\sigma_1\sin\alpha_1}\tan^{-1}\left(\frac{t_1\rho\cos\sigma_1}{1-t_1\rho\sin\sigma_1}\right),$$

$$\theta_2 = 2\tan^{-1}\frac{\rho\cos\sigma_0}{1+\rho\sin\sigma_0} - \frac{2\cos\sigma_0}{\cos\sigma_1\sin\alpha_1}\tan^{-1}\left(\frac{t_1\rho\cos\sigma_1}{1+t_1\rho\sin\sigma_1}\right);$$

and

$$\tfrac{1}{2}(x_{\rm A}-x_{\rm A'}) = -\sin\sigma_0\left(1+\frac{\pi}{2}\cos\sigma_0\right)\left(\frac{\cos\alpha_1}{1+\cos\alpha_1}\right)^2$$

$$-\int_1^0\left\{\frac{1-\rho^2}{4\rho^3}\left[(1-2\rho\sin\sigma_0+\rho^2)\cos\theta_1-(1+2\rho\sin\sigma_0+\rho^2)\cos\theta_2\right]+\frac{\sin\sigma_0}{\rho^2}\right\}d\rho,$$

$$\tfrac{1}{2}(y_{\rm A}-y_{\rm A'}) = \cos\sigma_0\left(1+\frac{\pi}{2}\cos\sigma_0\right)\left(\frac{\cos\alpha_1}{1+\cos\alpha_1}\right)^2$$

$$-\int_0^1\left\{\frac{1-\rho^2}{4\rho^3}\left[(1-2\rho\sin\sigma_0+\rho^2)\sin\theta_1+(1+2\rho\sin\sigma_0+\rho^2)\sin\theta_2\right]\right.$$

$$\left.-\frac{\cos\sigma_0}{\rho^2}\left(\frac{\cos\alpha_1}{1+\cos\alpha_1}\right)\right\}d\rho.$$

These give the position of the origin referred to the point D, and also the distance AA′ as a verification of the barrier itself, when plotted by means of the values of R, θ in terms of σ.

Address of the President, Sir Charles S. Sherrington, at the Anniversary Meeting, November 30, 1922.

The list of deaths from the Society's roll for the just completed year is somewhat exceptionally large. Early last December, within two days' interval, died HARDINGE GIFFARD, Earl of HALSBURY, and NATHANIEL, Lord LINDLEY, veterans of the Judicature, elected to the Society in 1888 and 1898 respectively. Lord Lindley was son of a former Fellow, the Botanist, John Lindley.

Later in December died THOMAS ALGERNON CHAPMAN, entomologist. From 1868 onwards, even to this present year, valuable papers from him had appeared, chiefly on the Lepidoptera in their bionomic and taxonomic bearings.

Also in December died GEORGE STEWARDSON BRADY, for many years Professor of Natural History in the College of Medicine, Newcastle-on-Tyne. His work dealt mainly with the Crustacea; the "Challenger" collection of Copepoda and Ostracoda were described by him.

Passing now to the present calendar year: JAMES COTTERILL, sometime Professor of Applied Mathematics in the Royal Naval College, and previously at the Royal School of Naval Architecture before its removal from South Kensington.

JOHN KIRK, member of the Civil Medical Staff during the Crimean War, then physician, naturalist and chief officer to Livingstone's second expedition, thereafter for a score of years trusted representative of this country in Zanzibar, and later, after his return to England, frequently appealed to by the Foreign Office as a valued authority on African questions. A gifted naturalist, his identification of *Strophanthus* as the source of an African arrow-poison conduced to the introduction by pharmacology of one of its valued modern remedies. He made large contributions to economic botany. Collections received from him greatly enriched the Natural History Museum and the Kew Herbarium. The experimental plantation, established at his expense, near Zanzibar, supplied observations of the highest value. Important Botanical and Zoological species bear his name, thus, *Sanseveria Kirkii Landolphia Kirkii* and *Madoqua Kirkii*, Kirk's Gazelle. The Kirk Range west of the Shire River, was named by Livingstone after him. It has been well said that these and like tributes to his name fitly memorialise it among settlers in lands secured for them by his far-seeing statesmancraft.

JAMES, Viscount BRYCE, distinguished writer and public servant, amid other ceaseless preoccupations, always an observer and lover of the face of Nature.

William Henry Mahoney Christie, Astronomer Royal for nearly thirty years. In January last, after starting on a sea voyage, he died on shipboard before reaching Gibraltar. His period of office at Greenwich was notable for large additions to the Equipment of the Observatory and the enablement of Greenwich to take its part in the photographic mapping of the heavens.

Benjamin Moore, biochemist, dying little more than a year after appointment to the Whitley Chair at Oxford. He was joint founder with Mr. Edward Whitley, of the 'Biochemical Journal.' His scientific interest lay especially in the applications of physical chemistry to biology. He accomplished much research and inspired much research in younger men. His work was diverse and fresh of view. Of an imaginative and sensitive nature, occasionally somewhat impatient of critical control, he chose bold themes and handled them boldly. He rendered great service to the Medical School of Liverpool University as its Dean through a number of eventful years in its upgrowth. A facile speaker and writer, he took active part in various movements for new public measures towards industrial hygiene.

Augustus Desiré Waller, died in his sixty-sixth year after but few days of illness. A physiologist, as had been his father Augustus, likewise a Fellow of the Society. Waller will be remembered perhaps chiefly as the first to show that the electric currents set up by the beating of the heart can be recorded by leads from the surface of the human body. His discovery was the more remarkable because made in the eighties, long before the introduction of the string galvanometer. His method after a time developed into the clinical electrocardiography of to-day. He accomplished much other work as well. Latterly the so-called psychogalvanic reaction accompanying transient emotion had attracted him; with it dealt several of his latest papers to the Society. For the fifteen years immediately preceding his death he had been in charge of the London University Laboratory of Physiology, at South Kensington, a laboratory which owed its existence largely to Waller's own initiation and support. Physiological experimentation was at once his profession and his hobby, occupying his professional hours, his family life, and his holidays no less. He was a lucid and incisive talker and writer. Born in Paris, his circle of scientific friends and acquaintance was almost as numerously French as English; he formed in several ways a living liaison between the French and British physiological schools.

Marie Camille Jordan, a Foreign Member of the Society, held the Chair of Mathematical Analysis at the Ecole Polytechnique, Paris, and was editor of the 'Journal de Mathématique' ('Liouville's Journal'). His earlier work lay in the theory of groups, and the results of his researches are embodied in his 'Traité des Substitutions,' where he obtained fundamental results on primitive transitive and composite groups and on the composition-factors of a group. To him is owing, for example, the concept of a function of bounded variation. But his best known work was probably in the region of analysis situs; it was he who first gave a precise analytical account of the geometrical notion of a curve; and 'Jordan's theorem,' that a closed Jordan curve divides the plane into two distinct connected regions, is one of the most famous theorems of modern mathematics. Much of Jordan's later work is contained in the later editions of his classical 'Traité d'Analyse,' a work which has played a great part in the education of most of the prominent mathematicians of our day.

GEORGE BALLARD MATHEWS, mathematician died in his sixtieth year, after prolonged illness. His original work lay chiefly in contribution to the theory of ideal numbers. On this subject he was author of the only serious English Treatise, a work of which, unhappily, one volume only has been completed. He was also an accomplished analyst and geometer, and a critic of acuteness and range.

PATRICK MANSON left a name long to be remembered for pioneer work and discovery in the field of the parasitology of tropical disease. He tracked the mosquito as the carrier in filariasis, and suggested it also as likely for malaria. His paper entitled "The Metamorphosis of *Filaria Sanguinis Hominis* in the Mosquito," appeared in the Linnean Society's 'Transactions' as far back as 1884. Returning from China, the scene of his earlier professional years, in 1898, he succeeded, during the Colonial Secretaryship of Mr. Joseph Chamberlain, in getting started the London School of Tropical Medicine. The Manson Memorial Medal, this year instituted there for triennial award to work of special distinction in Tropical Medicine, is a tribute to Manson's work of example and leadership in that field of medical science.

ALFRED BRAY KEMPE was a mathematician of distinction; in the intervals left between less abstract pursuits he made himself an acknowledged authority on the Theory of Linkages, a branch of kinematical geometry, and on the application of Mathematics to Formal Logic. He was President of the London Mathematical Society from 1894 to 1896. When later, legal duties as Diocesan Chancellor preoccupied his attention, he yet found time to render good service to science as Treasurer to this Society from 1898 until 1919. By virtue of his office as Treasurer to the Society, he was also Treasurer of the National Physical Laboratory from its formation until its transference to the Government in April, 1918. His time and thought were always unstintingly given to the responsibilities of these offices. Universally esteemed and respected, he will be gratefully remembered for the labour and careful attention and advice which the Society had from him constantly throughout more than twenty years.

CHARLES LOUIS ALFONSE LAVERAN, Foreign Member, died in Paris, where for much of his later life he had been on the staff of the Institut Pasteur. It was while a Military Surgeon in Algiers from 1878 to 1883 that he had commenced his studies of Malaria. He was the first to describe and recognise as such the malarial parasite, detecting and calling attention to certain minute amœboid bodies containing pigment and present in the circulating blood. This may be regarded as the initial step in that course of discoveries due to successive observers, more than one of them belonging to this Society, which has resulted in the remarkable degree of scientific control of malaria now happily possible. Laveran was Nobel Laureate in Medicine in 1907, and was elected into our Foreign List in 1916. As recently as 1917 he issued his comprehensive treatise "Leishmaniases." Retiring and reticent in manner, he continued into advanced life the even tenour of his systematic industry, although his most outstanding discovery now dates from more than 40 years ago.

On June 4 the death of WILLIAM HALSE RIVERS-RIVERS cut short in the fulness of his activity and powers a psychologist and ethnologist of exceptional significance. Indefatigable in research, he had opened up new lines of enquiry. His thoroughness not rarely so lit side issues that they were seen to be main ones. Thus, a table of blood-descent, worked out for a closed island community in the course of colour-vision observations in Oceania, opened his kinship method employed later with striking success in his "History of Melanesian Society." Again, in the research with Dr. Head into skin sensations, conjoint originality and freshness of view marked their investigation as a re-start in the whole subject. Rivers brought to ethnology and psychology along with the training of a skilled physician the observational flair of a born naturalist. For his career ethnology and psychology in a sense competed. As to ethnology he felt strongly that problems were urgent by reason of the vanishing of primitive societies under alien influences so that the data if not collected now could never be obtained. After starting, therefore, as a laboratory psychologist he had later, partly impelled by public-spirited motive, become chiefly a sociologist of primitive peoples; he sojourned to that end at various times in Upper Egypt, in Melanesia, and Southern India. But with the outbreak of the war a greater urgency recalled him to psychology. Arriving home from his second visit to Melanesia, he was taken by his commission in the R.A.M.C. to the hospital ward, there to deal with the psychoneuroses flooding in from the war fronts. He was among the earliest to understand those cases. He greatly contributed to the recognition of their mode of causation and to the adoption of rational treatment for them. He gave to them the very best that was in him, and in so doing he helped to open a new chapter in scientific psychology both for himself and for others. His volume, "Instinct and the Unconscious," first appearing in 1920, is already in its second edition; another almost completed work, "Conflict and Dream," is to be published posthumously. Rivers, when death took him, was in the exceptional scientific position of being regarded as a leader in intellectual enterprises largely of his own creation in at least two separate though sister branches of science. It is of interest that, highly trained in introspection as he was, he has left a recently written comment on his own method of thought.

"I," he wrote, "am one of those persons whose normal waking life is almost wholly free from sensory imagery, either visual, auditory, tactile, or any other kind. Through the experience of dreams, of half-waking, half-sleeping state, and of slight delirium in fever, I am quite familiar with imagery, especially of a visual kind, which, so far as I can tell, corresponds with that of the normal experience of others. I am able to recognise also that in the fully waking state I have imagery of the same order, but in general it is so faint and fragmentary that the closest scrutiny is required for its detection. It is clear to me that if it were not for my special knowledge and interest I should be wholly ignorant of its existence."

Among Rivers' characteristics were great and methodical industry, and a breadth of view which detail, however complex, never obscured. Apart from the reticence of modesty he was, as perhaps the above quotation may have served to show, engagingly frank; he was also extremely generous and was entirely free from all fear except that of missing the truth.

WILLIAM CARRUTHERS, botanist, for 24 years Keeper in the Department of Botany at the British Museum: his tenure included the time of removal of the collections to South Kensington, and many consequent expansions and improvements. He had been a Fellow of the Society for more than 50 years.

WILLIAM GOWLAND, sometime Professor of Metallurgy at the Royal School of Mines, a recognised authority on the non-ferrous metals. He had been in earlier years assayer to the Imperial Japanese Mint at Osaka. A man of many parts, his variety of attainment is witnessed by the circumstances that he presided—in fact in one and the same year—both at the Institution of Mining and Metallurgy and at the Royal Anthropological Institute.

JACOBUS CORNELIUS KAPTEYN, astronomer, elected Foreign Member in 1919. Early appointed Professor of Astronomy at Groningen, where, possessing no observatory, he turned to computational work, and undertook the measuring and reductions of photographs of the Southern Stellar Survey, sent him by Sir David Gill. This and similar work he carried on devotedly through many years. During his studies on proper motion he made the notable discovery of the two star-drifts. His latest paper on the configuration and motion of the Stellar System was published a few days before his death.

ARTHUR RANSOME, physician, died in his eighty-eighth year after long retirement from practice. He was the author of a number of papers on the etiology of Tuberculosis.

JAMES ARTHUR POLLOCK, Professor of Physics in the University of Sydney since 1899, died at the age of 57. A versatile and extremely skilful experimenter. Some of his measurements of specific inductive capacity can claim to be the most exact and trustworthy extant. He also carried out an important investigation of the ions of the atmosphere. He, as though he had been a man of younger years, entered for active service, and did valued work in range-finding on the Flanders front.

DAVID SHARP, entomologist, author of many papers, especially on the Coleoptera. He was for some years in medical practice. For nineteen years he was Curator of the Insect Collections at Cambridge, and contributed the two volumes on the Insecta to the Cambridge Natural History. He was general Editor of the 'Zoological Record' from 1892 onwards into the present year. Throughout that time he was himself recorder of the whole literature on Insects.

FREDERICK THOMAS TROUTON, physicist, a pupil at Trinity College, Dublin, of FitzGerald, and later his colleague there. He contributed toward the proof by experiment of the common electro-magnetic origin of ordinary light and of Hertzian waves. His researches were continued at University College, London, on his becoming Quain Professor there, into the effect of charging a condenser moving in the plane of its plates; and on the influence on the electrical resistance of a wire of moving it in and across the æther stream. For the last ten years of his life his activity as an experimentalist had been sadly crippled by illness.

EDMOND HERBERT GROVE-HILLS died on the 2nd of last month in his fifty-ninth year. Winchester, Woolwich, a Commission in the Royal Engineers, distinguished work in the Topographical Section of the General Staff, and Brigadier-General in 1918; these were phases in a life consistent in its pursuit of science, and especially of geodesy and astronomy. He was President of the Royal Astronomical Society from 1913 to 1915. Of great natural ability, with gifts of organising and administration, he will be missed from this Society, in whose aims and welfare he took constant and public spirited interest, and among whose members he had not a few close friendships.

CARGILL GILSTON KNOTT had been Professor of Physics in the University of Tokio, and had early helped forward in Japan the study of magnetics and seismology. Returning to Edinburgh, he then became Reader in Applied Mathematics in the University, where as a student he had been under, and later acted as Assistant, to Prof. Tait. Of his writings, perhaps that best known is his "Biography of Tait." For the last ten years of his life he had rendered valued service to the Royal Society of Edinburgh as its General Secretary.

ALEXANDER CRUM BROWN, died in his eighty-fifth year. Emeritus Professor of Chemistry in the University of Edinburgh, he had retired in 1908 from the Chair to which, in 1869, he had been elected in succession to Lyon Playfair. Versatile in and of varied accomplishments, his interests followed literary equally with scientific pursuits.

HENRY JOHN ELWES, naturalist and traveller; a collector for whom no journey was too arduous. His book "The Trees of Great Britain," is a monumental work.

ISAAC BAYLEY BALFOUR, botanist, as was his father, who likewise was a Fellow of this Society. Early distinguished in systematic botany he became Professor successively in the Universities of Glasgow, Oxford, and Edinburgh, and concurrently with his Edinburgh Chair, King's Botanist for Scotland, and Regius Keeper of the Royal Botanic Gardens, Edinburgh. His special researches lay chiefly with rhododendrons and primulas, particularly their recently found Chinese species. His whole career was fruitful for advance both of Botany and Horticulture. An enthusiastic nature, gifted with lucid expression and practical wisdom, was in Balfour devoted unreservedly to the service of the science with which, for two generations, his name has been conspicuously associated.

Some here present may happen to recollect that Sir Alfred Kempe, when Treasurer, was wont to comment on the longevity attaching to the F.R.S. Of the Fellows of the Society who have passed away during the just completed twelvemonth no fewer than seven had exceeded their eighty-seventh year.

Turning now to other matters, it will be remembered that last year the Society received by bequest a benefaction from the late Miss L. A. Foulerton, who by gift had already founded the Foulerton Studentship. For following out the wishes of the testatrix in regard to her bequest, the decision come to, on the advice of a large and representative committee dealing with the matter, has been that by means of the benefaction there should be created "one or more Research Professorships." These Professorships must, by the terms of the bequest, be assigned to subjects coming within the range of the defined though extensive field of science specified in the bequest. But the step thus taken and the precedent thus set seem sufficiently important and novel to be of interest to the Society's whole circle.

In the words of its Charter, repeated at the admission of each new Fellow, the Society is described as instituted "for Improving Natural Knowledge." A main means for that "improving" is discovery. In the case of Natural Knowledge the main road to discovery must lie in research. There are several ways in which research can be encouraged, and one of them lies in providing to suitable workers the means enabling them to devote their time freely to investigation. The Society is fortunate in possessing now to a somewhat greater extent than formerly funds that may be considered as permanently allocated to this fundamental object. In addition to the Research Professorship whose tenure is about to begin, it has five Research Studentships, namely, the Mackinnon, Sorby, Tyndall, Moseley, and Foulerton. All of these are of recent foundation; the oldest, the Mackinnon, does not date back yet to a full quarter of a century. The Society's existence extends now to more than two and a half centuries; financial help directed to this eminently important aim has, however, come only relatively recently. That it should have now begun may be a sign of the arrival of an Age in some respects new; the beginning of a trend towards wider public interest in and sympathy with research. Such a reflection turns our eyes naturally to the progress of endowment of research which has now for some time become evident and familiar when we regard America. The Rockefeller Foundation exemplifies munificence in that direction. The liberality with which private benefaction across the Atlantic has aided and is aiding research undertakings is too striking to need insistence on. I would only remark as to it that the set of its current, far from showing any ebb, flows more and more. This argues that here too, once begun, it will not decline. A movement that justifies itself should increase and extend. It behoves us, therefore, at this early stage of what may hopefully be regarded as a movement, to view as a whole our existing opportunities, relatively small though as yet they be. Arisen desultorily and recently, our endowments for this purpose, namely, the maintenance or partial maintenance of research workers, have yet arranged themselves into something of a scheme. The Studentships and the one or more Professorships form together a series, reaching at the one end from opportunities for workers of promise to carry their career toward fulfilment of that early promise to, at the other end, provision for men of already proved achievement to devote themselves unrestrictedly to the full pursuance of investigation wherein they have made their mark.

A salient feature in the administration by the Society of these emoluments to investigators is, that they are in no way tied up for award to any particular

Laboratory, University or Institution. The Society has no laboratory of its own, and the Society's administration of the emoluments takes advantage of the facilities of existing research institutions, University or other, and endows its appropriately selected workers wherever in those institutions they may happen to be, presupposing always the existence there of due facilities for their researches. Our research emoluments, therefore, do not compete with or tend to deplete existing research institutions, but to co-operate with and reinforce them. They do not withdraw the researcher from the University or other research Institute to which he already belongs. This dovetailing in with existent research appears especially important in its regard to the Universities, because it affords help, needed by Universities, for the execution of research; and it seems a happy feature of the scheme that it assists research where that can be prosecuted in contact, so to say, with students, thus tending to promote research schools.

And while on the subject of co-operation as a factor in the promotion of research I would advert for a moment to wider co-operation in that aim. The Rockefeller Foundation mentioned just now supplies an instance. Breadth of outlook and sympathy truly noble characterizes the distribution of the Rockefeller donations to research. Privately endowed Corporation as it is, its disbursements cross national frontiers with a liberality that sets a high example of the combination of nationhoods in the cause of scientific discovery. In London, University College had a great Rockefeller gift last year. In the present year has come from the Foundation promise of a larger gift to the British Government for the establishment of a School of Hygiene in London, for the better recruitment and development of the services of Preventive Medicine not only in this country itself but throughout the Empire. And again in this present year in Montreal, a great gift is extended from the Rockefeller Foundation to Science and Medicine at McGill University in memory of the McGill's fallen in the war. The Foundation, by its courtesy and hospitality, has been the means, not once but often, of personal opportunity for exchange of scientific ideas and experience between investigators in America and Europe. There can be in the history of research benefactions none of better augury for the widening of international co-operation than that of the Rockefeller Foundation. And turning to the Society's own experience in the present year we have an admirable instance of wide co-operation in the High Altitude Expedition to Peru, under Mr. Barcroft's leadership. For that expedition, to the great advantage of the undertaking, there co-operated along with this Society, both as regards men and money, Toronto University and the Harvard Medical School, along with Cambridge and Edinburgh.

And it is pleasing in reviewing the past year to note the international mark of appreciation which has come to two of the Fellows of the Society, in the award of the Nobel Prizes for Physics and Chemistry to Dr. Aston and to Prof. Soddy respectively.

Finally, of events in Biological Science in the past year I may mention one that is attracting attention at this time. In the Physiological Laboratory of Toronto University has been prepared a pancreatic extract possessing striking power over the carbohydrate metabolism of the body. Potent as it is, experience with it is still limited. Work of urgency is required with what may prove to be a desired remedy; the first programme is further investigation of the extract's full properties, with caution as to raising hopes which practice may but partly fulfil. Such are the considerations which weigh with the Canadian—and the discovery is Canadian—University to whom the discovery is due. In this country the Medical Research Council has undertaken public-spirited direction of the extract's preparation and of further determination of its properties.

The physiological steps of the discovery may be briefly outlined thus:—

Destruction of the pancreas is well known to produce in the dog a diabetes-like condition, rapidly fatal. The liver's store of glycogen is lost, and cannot be renewed by even liberal supply of its normal source, carbohydrate food. Sugar formation from proteins ensues, with rapid wasting of the tissues; at the same time the blood is surcharged with sugar, and the tissues are unable to make use of sugar. In a normal animal, glucose put into the circulation raises the ratio of CO_2 expired to Oxygen absorbed, because the tissues consume the sugar. But glucose similarly introduced into the depancreated diabetic animal does not raise the respiratory quotient; the tissues no longer consume the sugar.

The inference has long been that the pancreas produces some substance enabling the body to make use of sugar—some substance that in fact should control certain forms of diabetes. At Toronto there seems to have been secured the extraction of that substance.

The pancreas consists of two structures intimately commingled. The one, secreting cells set round ducts into which they pour the pancreatic juice, potently digestive. The other, scattered in tiny islets seemingly unrelated to the ducts though closely related to the blood channels. The want of success of pancreatic extracts in mitigating a diabetic condition might be due to digestive powers of the juice cells destroying an anti-diabetic substance of the islet-cells. Dr. F. C. Banting, determined to avoid this possibility by preparing extracts made from the pancreas after its trypsin-yielding cells had been selectively brought to atrophy by ligation of the gland ducts.

He and Mr. Best, a collaborator who joined him, overcoming formidable difficulties of technique, succeeded in preparing the required material, and in examining the effect of extract upon diabetic depancreated dogs. They found the sugar fall both in the blood and urine, and that the animals, instead of dying in three weeks, remained, while treated, in excellent condition. The further prosecution of the work subsequently engaged other collaborators: to mention them in alphabetical order, Collip, Hepburn, Latchford, Macleod and Noble; of these Prof. Macleod, himself Director of the Toronto Physiological Laboratory, is well known as a skilled authority in experiments on carbohydrate metabolism, and Dr. Collip is Professor of Bio-Chemistry in the University of Alberta, though temporarily working at Toronto. With team work, advance has proceeded relatively quickly, and successful extracts are now obtained from ordinary ox and other pancreas.

Of much physiological interest is the fact that the active principle in the extract seems one *normally* controlling the blood-sugar in health, its injection rapidly lessens the blood-sugar in normal animals.

The extract, added to a simple perfusion fluid containing a little glucose and streamed through the isolated rabbit heart, increases three- or fourfold the heart's uptake of sugar from the fluid. The extract sometimes evokes serious nervous disturbances seemingly associated with extreme fall in the amount of the blood-sugar.

Administered to diabetic depancreated animals, the extract brings reappearance of the liver's glycogen store, while bringing down the sugar excess in the blood and the excretion of sugar and acetone in the urine; and it enables the diabetic organism to consume sugar. It also lessens or prevents hyperglycæmia produced in animals in several other ways.

Gratifying success has already attended the use of this extract in the relief of diabetic patients; much further research is, however, yet needed for development of the methods of extraction and of the routine use of the active principle.

The important physiological advance thus just reached comes as a fit reward to those who have achieved it. It is of course as the striking crowning of steady work pursued by many various workers through many earlier years Such work, we may remember, lay often open to charge by the unenlightened of being merely academic and fruitless; its reward being at the time simply the intrinsic scientific interest of the facts obtained. The Toronto investigators we may be sure would say with Pasteur, "To have the fruit there must have been cultivation of the tree." Part of the merit of the recent successful investigation has been its appreciation of possibilities indicated by previous work. But that merit is after all only a preliminary to the main achievement. The actual achievement is the deserved success of a bold attack conducted with conviction and determination, and carried through in the face of formidable experimental difficulties. High praise is due to Dr. Banting and Mr. Best and to their subsequent collaborators.

Finally, I must not conclude without mention of two events which have happened only to-day. This afternoon the first appointment to the Foulerton Professorship has been made. Prof. Starling has been elected to it; an election of happy augury for the success of the Professorship. The fine work he has done is well known not only to us of this Society but to Physiology and Medicine throughout the world. As first Foulerton Professor he is the distinguished forerunner of what we may hope will be a long line of distinguished successors in the Foulerton Chair.

Also this afternoon the Society has received at the hands of Sir Richard Threlfall and Dr. Scott on behalf of the subscribers the gift of a portrait of Sir Joseph Thomson. The portrait is by Mr. Fiddes Watt, an admirable work of art and a striking likeness. In the Society's rooms it will be a cherished memento of one whose scientific achievements are a source of pride to the whole Society, of one whose services as President and in other capacities the Society holds in grateful remembrance, and of one whose presence at the Society, always valued as it is, the Society is glad to have this afternoon and hopes will long continue to be frequent. The portrait will typify on our walls much of the very best of the Society's tradition.

I now proceed to the presentation of the Medals.

The Copley Medal is awarded to Sir Ernest Rutherford.

Sir Ernest Rutherford early turned his attention to the ionising radiations discovered by Becquerel. He showed that these rays could be divided into three classes, α, β and γ, according to their penetrative and other properties. The β rays were rapidly moving electrons. He showed that the α rays were much heavier, and subsequently he proved them to be helium atoms with a positive charge equal to twice that of the electron. The γ rays were of electro-magnetic type and were not deflected by a magnetic field.

With characteristic relentlessness he investigated irregularities noted in the ionisation produced by thorium compounds, and was led to the discovery of the induced activity and the emanations. The latter he proved to be radio active gases. A study of the time-changes of these brought him in collaboration with Prof. Soddy to put forward the theory of natural atomic transmutations, confirmed experimentally in various ways. More recently, he and his pupils have been especially concerned with the deflections of α particles in their passage through matter. In result of his experiments he was led to the view that the positive electric charge in the atom is confined to a minute nuclear region in the atom, that that region comprises nearly the whole mass of the atom and has a charge equal to the electronic charge multiplied by the atomic number of the element.

In this work the *a* particles were located by the scintillations which they produced on a zinc sulphide screen. It was found that when the screen was beyond reach of the original *a* particles a number, relatively small, of scintillations still remained. By careful analysis of this phenomenon he has shown that in some cases these additional effects must be due to hydrogen atoms ejected from the nuclei of the different elements by the bombarding *a* particles, and that this disruption takes place at expense of energy latent in the disrupting atom. Thus he has passed on from discovering the natural transmutation of the elements to producing and successfully observing the artificial transmutation of them.

The Rumford Medal is awarded to Prof. Pieter Zeeman.

Prof. Zeeman's discovery of the splitting up of spectroscopic lines under the influence of magnetic force had important results, not only owing to its theoretical consequences, but also in its applications to Celestial Physics, enabling astronomers to trace magnetic effects at the surface of the sun. Among Zeeman's subsequent contributions to science, special reference may be made to a recent investigation dealing with the propagation of light in moving bodies. In all earlier experiments the dispersion of light in the medium was neglected, and the irregularities in the flow of the liquid, constituting the moving body, prevented accurate measurements. Zeeman set himself the task of obtaining greater accuracy, and for this purpose to investigate the effects in solid substances such as quartz or glass, giving to these bodies an oscillatory velocity, and applying an instantaneous photographic method, the exposure taking place when the velocity was at its maximum. Both the optical and mechanical arrangements presented considerable difficulties requiring all the skill of an accomplished experimenter.

A Royal Medal is awarded to Mr. Joseph Barcroft.

For the last twenty years Mr. Barcroft has been prominent among physiologists, owing to his researches on the respiratory function of the blood and its relation to the activity of the tissues. He has with various collaborators worked out the changes in the normal consumption of oxygen accompanying functional activity in various representative organs—salivary gland, kidney, cardiac and skeletal muscle and liver. He has also worked out and thrown new light on the meaning of the dissociation curve for oxygen.

exhibited by blood and by pure hæmoglobin, and on the influence of dissolved salts upon that curve. In prosecution of research on respiratory he has explored the conditions of life at high altitudes, himself being leader of two expeditions for that purpose, one to Monte Rosa and a second late last year to the high Andes, bringing back this year observations of decisive importance. His work is of fundamental character and high precision and a number of his methods have now passed into world-wide use.

A Royal Medal is awarded to Mr. Charles Thomas Rees Wilson.

Previous work having shown the important part played by dust particles in the condensation of super-saturated vapour, to Mr. Wilson is due the demonstration that the ions produced by the passages of X-ray act in a similar manner, thus giving a palpable proof of the discrete nature of the ions apart from their electrical effects. Later, he was able on the same principles to render visible, and to photograph, the actual path of an *a* particle through a gas, as indicated by the ions which it produces acting as condensation nuclei. More recently he has devoted himself to the study of the complex phenomena of atmospheric electricity. By methods which, though simple in principle, demand great skill in execution, he has not only measured the surface electrification of the ground, and thence the potential gradient, at any moment, but has also recorded its variation from instant to instant. Observations of this kind made during the progress of thunderstorms have enabled him to estimate the amount of electricity passing in a lightning flash. Wilson's work has been distinguished throughout by great experimental skill, whilst his writings are remarkable for their lucid exposition as well as by their sobriety of speculation.

The Davy Medal is awarded to Prof. Jocelyn Field Thorpe.

Prof. Thorpe commenced to publish in 1895, when the constitution of camphoric acid was attracting much attention among organic chemists. From 1895 to 1906 appeared a series of papers by Prof. Thorpe, on the synthesis of various substituted glutaric acids, and a variety of degradation products of camphoric acid. Perhaps the most striking of these researches were (*a*) the synthesis of isolaurolonic acid and a campholytic acid, and (*b*) a synthesis of camphoric acid itself.

The peculiar possibilities of ethyl cyanoacetate as a synthetic agent had been noted in these and other researches, and Thorpe investigated this substance more fully. As a result there have appeared an illuminating series of papers on the formation and reactions of imino compounds, giving rise to a variety of derivatives of naphthalene, hydrindene, pyridine, etc., and on the isomerism displayed by the glutaconic acids, for which an adequate explanation is now forthcoming.

In 1915 he published his paper on "Spiro Compounds," the first of a series dealing with the effect produced by the alteration of the tetrahedral angle, consequent on ring formation, on the formation and stability of a second ring joined to the existing ring by a quaternary carbon atom common to both. Experimental proofs have already been provided in support of the theory that an alteration in the angle subtended by two of the valencies of a carbon atom involves a corresponding alteration in that enclosed by the remaining two.

The Darwin Medal is awarded to Prof. Reginald Crundall Punnett.

Prof. Punnett has unravelled, with striking success, various extremely complex problems of inheritance. He was the first to find the true path in the interpretation of "coupling and repulsion," now termed "linkage." It was already known that at times factors belonging to distinct allelomorphic pairs were transmitted as if partially linked, but that also in other families the same factors might show repulsion.

Prof. Punnett conceived that these two phenomena must depend on parental combination, and this conception, subsequent joint experiments, with Dr. Bateson, confirmed. This discovery has led to very important extensions of genetic theory; most of the modern interpretations of sex-limited inheritance have grown out of it.

The above statement singles out as a convenient and striking example one of numerous studies due to him. He has contributed much other work also of importance to genetic science.

The Buchanan Medal is awarded to Sir David Bruce.

Trypanosoma Brucei, the causal organism of tsetse-fly disease is so named after its discoverer, Sir David Bruce, who likewise first showed its causal connection with that disease and with Nagana. The significance of this work for the whole study of the inter-relationship between human and animal disease of that kind was in itself a signal service to hygiene.

As Director of the Royal Society's Commission for investigating Sleeping Sickness, Bruce took a leading part in the elucidation of trypanosome infections, and in the adoption of counter measures against them. As Chairman of the Commission which traced the incidence in man of Mediterranean Fever to transmission through the milk of goats he contributed eminently to a great step of progress in the control of that disease.

During the war he presided over a Tetanus Committee and personally carried out the collection and analysis of data regarding Tetanus on a scale never previously attained. Later, as Chairman of the Committee on Trench Fever, he undertook the conduct of an enquiry which side by side with an investigation by our American allies, established in regard of that newly recognized disease, its transmission by lice and so indicated the means for its control.

The Sylvester Medal is awarded to Prof. Tullio Levi-Civita.

The investigations by Levi-Civita in pure geometry were the necessary foundations for the important physical discoveries of Einstein and Weyl.

Einstein, in the introduction to his most famous paper on " Relativity " writes: "The necessary mathematical apparatus for the general relativity theory lay ready to hand in the Absolute Differential Calculus . . . which was systematised by Rici and Levi-Civita." Weyl, in the introduction to his book containing his extensions of Einstein's theory to electro-magnetic phenomena, writes: " The discovery by Levi-Civita in 1917 of the conception of infinitesimal parallel-displacement gave me the impulse to a new examination of the mathematical foundation of Riemann's geometry." It is from this fundamental conception of parallel-displacement that Weyl develops his theory.

Levi-Civita, apart from these applications of certain parts of his researches, has shown himself one of the most fertile and original of investigators in differential geometry and theoretical mechanics.

The Hughes Medal is awarded to Dr. Francis William Aston.

Dr. Aston's researches on Isotopes, though begun only since the war, have already become classical. Starting from an observation by Sir J. J. Thomson that neon examined by the method of positive rays shows traces of a constituent of slightly higher atomic weight, Dr. Aston, by use of an ingenious method of focussing positive rays, has shown that a large number of the elements are complexes consisting of two or more kinds of atoms, having of course identical chemical properties but differing in atomic weight by one or more units. Except in the single instance of hydrogen, the atomic weight of each constituent is, to the limit of accuracy, a whole number on the basis of oxygen = 16. Thus the discrepancy of the experimental numbers from Prout's law in such cases as, for instance, chlorine, receives a simple explanation. Emphasis must be laid on the great manipulative skill required to achieve these results as well as on the high scientific importance of the results themselves.

On the Excitation of Characteristic X-Rays from Light Elements.

By J. C. McLennan, F.R.S., Professor of Physics, and Miss M. L. Clark, B.A., University of Toronto.

(Received May 3, 1922.)

I. INTRODUCTION.

In an attempt to fill up the gap between the shortest ultra-violet light waves hitherto produced and the longest X-ray waves known, Hughes* recently made a study of the characteristic X-rays emitted by carbon and by boron when bombarded by electrons. In this investigation the energy of the bombarding electrons was increased by steps, and the critical values were determined that were necessary and just sufficient to cause the bombarded element to emit its characteristic radiations with measurable intensities. These characteristic radiations were detected, and their intensities measured, by their photo-electric action on an insulated electrode of nickel or of silver.

The method followed by Hughes in recording his results was to plot curves with the values of the accelerating potentials of the electrons as abscissæ and the measures of the photo-electric effect divided by the corresponding electronic currents as ordinates. At certain critical accelerating voltages it was found that these curves showed marked and abrupt kinks or changes of curvature, and these changes were taken to connote the beginning of the emission by the bombarded element of its characteristic radiations.

By following this method he found two definite breaks in his curves for both carbon and boron, and these were taken by him to correspond to the critical K and L absorption wave-lengths for the two elements. For carbon, the breaks occurred at 215 volts and 34·5 volts, and for boron at 148 volts and 24·5 volts.

This would mean that the critical absorption wave-lengths of the K and L X-ray series for carbon were about $\lambda = 57{\cdot}5$ A. and $\lambda = 358$ A. respectively, and for boron about $\lambda = 83{\cdot}5$ A. and $\lambda = 505$ A. respectively.

Hughes' results appear to be the only ones as yet obtained with the element boron, but with carbon other researches have been carried out, and differing values have been found for the critical absorption wave-lengths of the K and L series for this element.

Kurth's† values are $\lambda = 42{\cdot}6$ A. (Acc. Pot. = 283 volts) and $\lambda = 375$ A.

---

* Hughes, 'Phil. Mag.,' p. 145 (January, 1922).

† Kurth, 'Phys. Rev.,' p. 461 (December, 1921).

(Acc. Pot. = 32·9 volts) respectively, Richardson and Bazzoni's* value for the K series is λ = 43·4 A. (Acc. Pot. = 286 volts), while Mohler and Foote's† results for the K series give λ = 45·4 A. (Acc. Pot. = 272 volts).

Hughes' value for the critical absorption wave-length of the K series of carbon, it will be seen, is considerably higher than any of the values found by the other investigators.

In an investigation recently carried out by us, in which we followed the method adopted by Hughes, the critical absorption wave-lengths of the K and L series were determined for the elements boron, beryllium and lithium and the critical absorption wave-length of the L series for carbon. The values found were as follows:—

Carbon, L series, λ = 166·7 A. (Acc. Pot. 74·0 volts).
Boron, K series, λ = 83·6 A. (Acc. Pot. 147·5 volts).
L series, λ = 292·2 A. (Acc. Pot. 42·2 volts).
Beryllium, K series, λ = 118·2 A. (Acc. Pot. 104·3 volts).
L series, λ = 428·1 A. (Acc. Pot. 28·8 volts).
Lithium, K series, λ = 290·8 A. (Acc. Pot. 42·4 volts).
L series, λ = 1019·0 A. (Acc. Pot. 12·1 volts).

For boron, it will be seen that our result for the K series is in good agreement with that obtained by Hughes. Our result for the L series critical absorption wave-length, however, differs considerably from that found by him. For beryllium and lithium, the values found appear to be the first as yet observed.

II. APPARATUS.

The apparatus used was made of pyrex glass and is shown in fig. 1. The bombarding electrons were obtained from a 9-mil. tungsten wire, HK, heated by a battery, A. The leading-in wires FH and KE were of negligible resistance and, as indicated, the connections were made to the battery so that there was a rise of potential in the filament from H to K. An ammeter, G, measured the strength of the heating current. This ammeter and the battery, A, together with the connecting and leading-in wires, were all carried on insulating supports of ebonite. The whole system HKEGDF was kept at any desired negative potential by means of the battery of small storage cells, C, attached at B. Since there was a rise in potential over the heated filament from H to K, the voltage of the battery, C, was taken to give a measure of the maximum value of the field between the filament and the anti-cathode, W. This anti-cathode was provided with a recess, L, into which the bombarded metal was pressed. Throughout the experiments the anti-cathode was maintained electrically at the potential of the earth and was kept at room temperature by means of a water-cooling system, arranged as shown in the diagram. The electrode, PR, was joined to a sensitive quadrant electrometer, and was terminated at P by a freshly-polished plate of nickel. T was a brass tube with the end at N capped by a gauze disc with meshes about 4 sq. mm. in area. X was an insulating extension tube of ebonite attached to T. It, also, was capped with a gauze disc shown at M. By means of the battery, O, which was insulated, this gauze, M, was always kept at a potential of about 10 volts below the lowest potential of any part of the filament, HK. The object in using it was to prevent diffusion of negative ions to the electrode, P. The tube, T, and therefore the gauze attached to it, N, was kept in all the experiments, by means of the battery, Q, at a positive potential of about 400 volts above that of the earth. The use of the tube, T, and the gauze, N, was twofold. Its charge provided a positive field for the photo-electric current from P, and it acted as an electric screen to prevent any diffusion to P of positive ions from the region about L and HK. The guard tube, S, made of brass, was kept joined to earth, and so prevented any leakage of electricity from the charged tube, T, and its gauze cap, N, to the electrode, PR.

---

* Richardson and Bazzoni, 'Phil. Mag.,' p. 1015 (December, 1921).

† Mohler and Foote, 'Scien. Papers, Bur. of Standards,' No. 425 (1922).

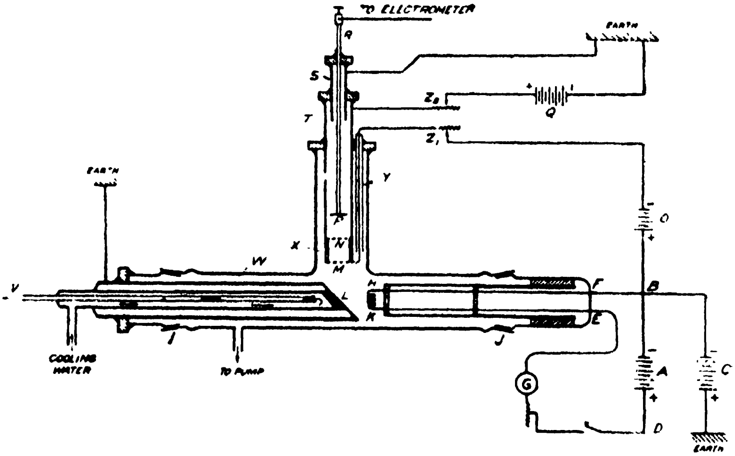


Fig. 1.

The apparatus was exhausted as highly as possible with a Langmuir mercury pump, backed by a Gaede mercury pump and a Gaede rotary oil pump.

The various parts of the tube were carried by ebonite insulating supports as shown in the diagram. All joints were made air-tight with sealing-wax. By means of the ground-glass joints at I and J, the anti-cathode, W, and the cathode, FEHK, could be readily withdrawn from, or inserted in, the apparatus, when desired. When the electrodes were placed in position the joints at I and J were coated with melted sealing-wax to make sure that they were air-tight.

III. METHOD OF EXPERIMENT.

In making an experiment the electron filament was first attached at H and K. The metal to be bombarded was melted into the recess in the anticathode at L, and the tube was sealed up and connected to the exhausting pumps. Electric connections were made as indicated, the electrode, R, being joined to the quadrants of the electrometer. Before any attempt was made to take readings the tube was generally evacuated for 10 or 15 hours, and at intervals the cathode, HK, was heated with current from the battery, A, and the metal in L bombarded with electrons accelerated by the field that the battery, C, provided.

After the long evacuation indicated, it was found that all gases were sufficiently extracted from the tube and the electrodes to enable one to take regular and consistent readings of the photo-electric current from the plate, P, under the stimulus of the radiation from L that fell upon it. When making the observations, care was taken to maintain the filament HK at a temperature sufficient to give a constant electronic current between it and L. The rise in potential from F to E was noted for all the readings. In carrying out our experiments, the potential of the system KHFBDG was lowered step by step by means of the battery, C, and for each step an observation was made on the photo-electric current from P. For low applied voltages the electrometer had to be used at its maximum sensitivity, but for high applied voltages it was necessary to reduce the sensitivity of the electrometer by attaching to its quadrants condensers of suitable capacity.

With all three of the metals, lithium, beryllium and boron, it was found that, when the exhaustion of the tube was carried sufficiently high, it was easy, by making slight variations in the current of the heating circuit, to maintain the electronic current from HK to L at a constant value for the periods required to take complete sets of readings over both the high and low ranges of applied voltages. With each metal between six and fifteen sets of readings were taken, and it was found that, apart from variations in the absolute values of the photo-electric currents due to the use of electronic currents of different intensities, all curves of the readings taken with each metal were of the same type and possessed the same characteristics as to their kinks or discontinuities.

IV. EXPERIMENTAL RESULTS.

(a) *Boron.*

Typical sets of readings taken for the higher range of applied potentials when boron was used as the anti-cathode are shown in Table I.

Table I.—Boron.

Higher range.				Lower range.	
Applied potentials.	Photo-elect. current (electrom. defln.).	Applied potentials.	Photo-elect. current (electrom. defln.).	Applied potentials.	Photo-elect., current (electrom. defln.).
volts	cm.	volts	cm.	volts	cm.
116·2	10·7	60·0	6·0	16·0	0·9
127·2	11·0	79·8	8·0	18·2	2·0
137·9	13·5	99·2	7·8	20·2	4·3
144·0	13·2	119·2	7·5	22·2	7·6
146·2	9·8	139·2	5·4	22·7	6·4
148·2	14·0	143·0	4·5	23·2	6·6
150·4	11·0	147·0	4·1	23·7	7·2
152·5	9·9	149·0	4·6	24·2	9·3
154·6	8·7	150·8	5·1	26·2	19·6
		154·8	5·8		
		156·8	6·1		
		157·8	5·6		
		177·8	4·7		
		197·8	4·5		
		217·8	4·3		
		254·0	3·8		

Curves representing these values are shown in figs. 2 and 3. These, it will be seen, exhibit discontinuities at 147 and 148 volts respectively. For the low range of applied voltages a typical set of readings is also given in

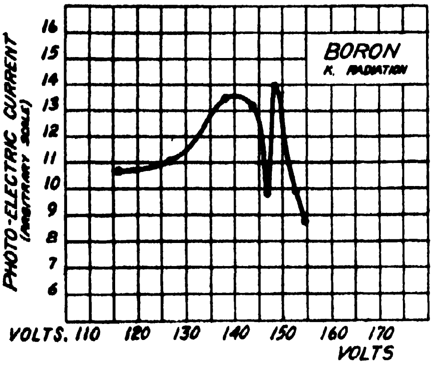


FIG. 2.

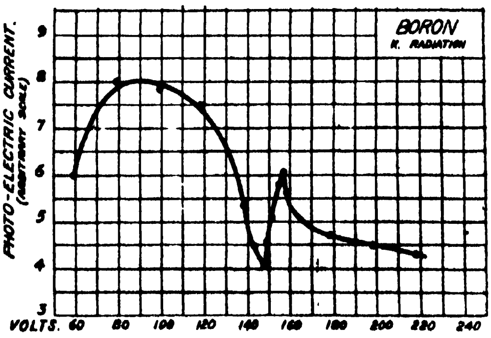


FIG. 3.

Table I, and a curve representing them is shown in fig. 4. In this case it will be seen that a marked discontinuity is indicated at 22·95 volts. In some of the curves for the low range of applied voltages with boron a second and less strongly marked discontinuity was observed at about 27·8 volts.

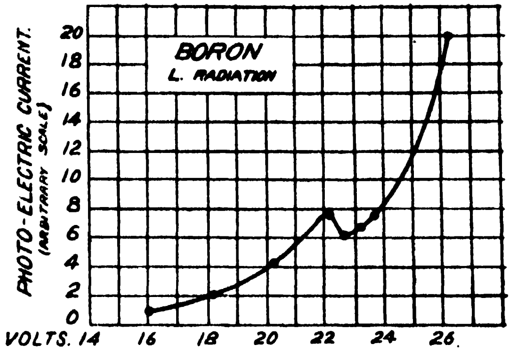


Fig. 4.

The discontinuities noted in the curves taken for boron are collected in Table II.

Table II.—Boron.

High range of potentials.			Low range of potentials.			
Curve.	Critical potential.	Voltage drop in filament.	Curve.	Critical potentials.		Voltage drop in filament.
				I.	II.	
	volts			volts	volts	
1	146	8·7	1	—	27·2	8·7
2	149	8·7	2	22·5	—	8·7
3	148	8·7	3	24·4	28·2	8·7
4	148	8·7	4	24·4	28·2	8·7
5	147	8·7	5	23·4	28·4	8·7
6	147	8·7	6	23·4	27·0	8·7
			7	24·2	28·5	8·7
			8	24·2	—	2·2
			9	23·0	—	2·67
			10	23·0	—	3·4
			11	23·45	—	2·67
			12	22·95	—	3·4
			13	22·95	—	3·76
			14	22·95	—	4·04
Mean =	147·5		Mean =	23·45	27·92	

From these it follows that a strongly marked discontinuity was obtained in the curves for boron at the mean voltage 147·5. A second strongly marked one was also obtained at the mean voltage 23·45, and a less strongly marked one at 27·92.

Table III.—Beryllium.

Higher range.				Lower range.	
Applied potential.	Photo-elect. current (electrom. defln.).	Applied potential.	Photo-elect. current (electrom. defln.).	Applied potential.	Photo-elect. current (electrom. defln.).
volts	cm.	volts	cm.	volts	cm.
39·6	4·05	38·8	4·6	11·2	5·9
50·0	4·2	49·6	4·8	13·5	7·6
60·7	4·7	59·8	5·7	14·0	6·1
71·0	5·3	70·2	7·4	14·5	4·6
73·2	5·6	72·4	9·2	15·0	4·5
74·8	4·9	73·8	8·2	15·5	4·2
77·0	4·9	76·0	7·4	16·0	4·1
79·0	5·5	78·0	7·6	16·5	4·6
80·6	5·9	79·6	8·6	17·0	5·0
95·7	6·8	94·3	11·0	17·6	10·1
106·3	10·0	104·6	18·5	18·0	15·2
116·1	14·4	113·8	29·0		
126·5	18·0				
137·2	21·5				

(b) *Beryllium.*

Typical results obtained with beryllium for both high and low applied voltages are shown in Table III. The corresponding curves are given in figs. 5, 6 and 7.

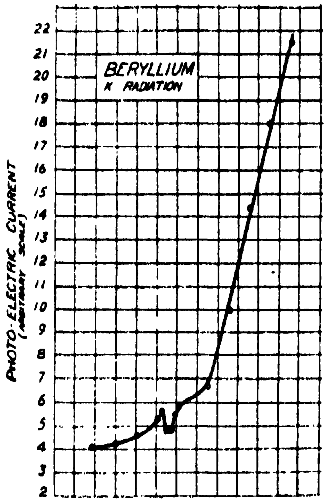


FIG. 5.

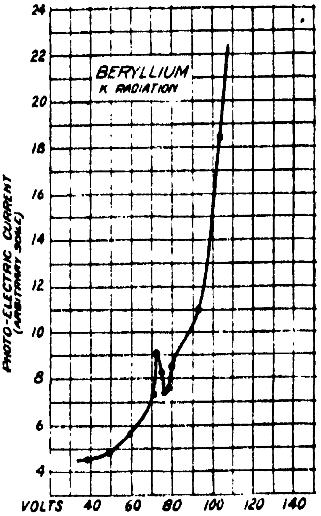


FIG. 6.

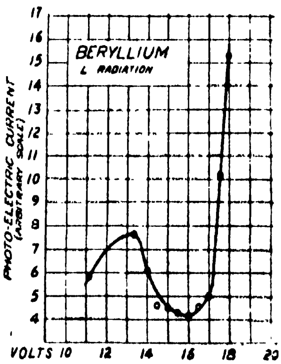


FIG. 7.

Table IV.—Beryllium.

Higher range.				Lower range.			
Curve.	Critical potential.		Voltage fall in filament.	Curve.	Critical potential.		Voltage fall in filament.
	I.	II.			I.	II.	
	volts	volts			volts	volts	
1	78·0	96	8·7	1	16·6	—	8·7
2	78·0	90	8·7	2	16·0	21·7	8·7
3	78·0	90	8·7	3	16·0	20·4	8·7
4	78·0	—	8·7	4	16·0	—	8·7
5	79·0	91	8·7	5	16·0	19·4	8·7
6	79·0	91	8·7	6	16·0	19·4	8·7
7	78·5	—	8·7	7	16·0	19·4	8·7
8	78·5	—	8·7	8	15·2	21·6	8·7
9	78·0	—	8·7	9	16·25	—	8·7
10	78·0	89	8·7	10	16·0	—	8·7
11	77·0	94	8·7	11	16·0	—	8·7
12	79·0	95	8·7				
13	78·5	—	8·7				
14	78·0	—	8·7				
15	78·4	97	8·7				
Mean =	78·2	93		Mean =	16·0	20·3	

Curve 5, it will be seen, shows discontinuities at 78·0 and 89·0 volts, and Curve 6 at 78 and 90 volts. Curve 7 shows a point of inflection, followed by a sharp rise at 16·0 volts. In some of the curves for the lower range of applied voltages a second and less strongly marked discontinuity was observed in the neighbourhood of 20 volts. With beryllium some fifteen sets of readings were taken for the higher range of voltages and eleven sets for the lower range. The critical voltages for all the curves are collected in Table IV.

The results given in Table IV for beryllium show that, in the higher range of applied voltages, a marked discontinuity was obtained on the average at 78·2 volts, and a less strongly marked one at 93 volts. For the lower range of voltages a definite discontinuity was observed for a mean voltage of 16·0, and, on the average, a less well defined one of 20·3 volts.

(c) *Lithium.*

Typical results obtained by us for lithium are collected in Table V, and curves representing them are shown in figs. 8, 9, and 10.

In Curve 8 there is a discontinuity at approximately 37·8 volts. All of the curves for lithium with the upper range of applied potentials exhibited kinks or discontinuities in the neighbourhood of 37 volts. Some,

Table V.—Lithium.

Higher range.				Lower range.	
Applied potential.	Photo-elect. current (electrom. defln.).	Applied potential.	Photo-elect. current (electrom. defln.).	Applied potential.	Photo-elect. current (electrom. defln.).
volts	cm.	volts	cm.	volts	cm.
28·0	5·0	26·0	1·8	1·6	0·5
30·0	7·5	27·6	3·3	4·2	1·0
32·0	14·0	29·6	4·1	5·6	3·0
34·0	15·7	31·8	5·3	7·6	3·6
36·0	13·5	33·8	8·6	9·0	3·8
38·2	33·5	36·0	13·0	11·0	4·0
39·8	58·7	38·0	21·5	12·8	4·9
41·8	84·4	40·0	39·0	14·8	6·4
44·0	113·9	42·0	101·4	16·8	7·9
		44·2	286·0	18·5	9·8
				20·0	11·8
				21·5	15·0
				28·6	18·8

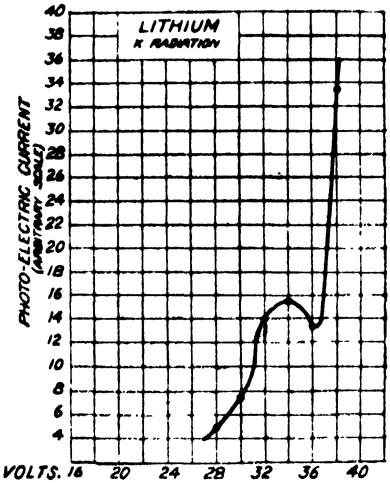


Fig. 8.

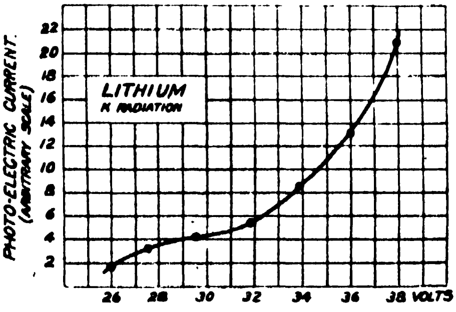


Fig. 9.

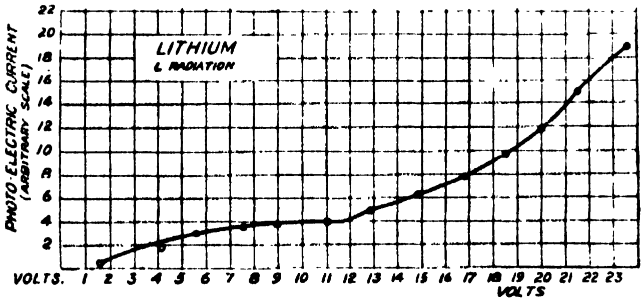


Fig. 10.

however, showed less well defined ones in the neighbourhood of 32 volts. The curve in fig. 9 is one of this latter type, a moderately well defined discontinuity being shown near 31·2 volts.

In the lower range of applied potentials a discontinuity was always observed in the curves with applied potentials in the neighbourhood of 12 volts. Such a discontinuity is shown in fig. 10. The critical voltages corresponding to the discontinuities noticed on all the curves for lithium are collected in Table VI.

Table VI.—Lithium.

Higher range.				Lower range.		
Curve.	Critical potential.		Voltage fall in filament.	Curve.	Critical potential.	Voltage fall in filament.
	I.	II.				
	volts	volts	volts		volts	volts
1	36·5	—	1·34	1	13·2	1·34
2	37·8	—	1·34	2	12·8	1·34
3	37·8	—	1·34	3	12·2	1·34
4	38·1	31·6	1·34	4	11·0	1·34
5	36·7	31·6	1·34	5	12·1	1·34
6	36·5	31·4	1·34	6	11·8	1·34
7	37·1	32·8	1·34	7	12·6	1·34
8	—	31·4	1·34	8	12·1	1·34
9	36·0	—	1·34	9	12·0	1·34
10	36·4	—	1·34	10	11·7	1·34
11	36·4	—	1·34	11	10·5	1·34
Mean	=37·0	=31·8		Mean	=12·0	

From these it will be seen there was a well-marked critical voltage for lithium with a mean value of 37·0, and a less definitely marked one with an average value of 31·8. In the lower range of applied potentials a critical voltage was always obtained, with its average value close to 12·0.

V. HIGH-RANGE CRITICAL VOLTAGES AND CRITICAL ABSORPTION K SERIES WAVE-LENGTHS.

An inherent defect in the method adopted in the present investigation lies in the fact that no means is provided by it for determining directly the significance of the observed critical voltages. Theoretical considerations, however, based on our knowledge of atomic structure, lead us to associate the well defined critical voltages that we observed in the higher range of applied potentials with the K X-ray spectral series of the three elements investigated.

Since the critical voltages 31·8 and 37 noted for lithium are in the ratio 27 to 32, it would appear that they represent the first two numbers of a series with a frequency formula given by $\nu = \mathrm{N}(1-1/m^2)$. Such a series would have for the critical potential of its shortest member 42·2 volts. From the relation $\mathrm{V} = h\nu/e = hc/e\lambda$, or $\lambda = 12331/\mathrm{V}$ (volts), it follows that the series would consist of a set of wave-lengths beginning approximately at 387·7 A. and extending to 290·8 A. Just how this series of wave-lengths originates is not clear, as no changes in state for the model of the lithium atom put forward by Bohr seem to make provision for it.

The critical voltages 78·2 and 93·0 observed by us with beryllium are also approximately in the ratio 27 to 32, and may also be taken to represent the first two members of a series with a frequency formula $\nu = \mathrm{N}(1-1/m^2)$. The critical potential corresponding to the last member of this series would be 104·3 volts.

The wave-lengths for the K series for beryllium, as calculated from the data given above, would extend from about $\lambda = 157{\cdot}6$ A. to $\lambda = 118{\cdot}2$ A.

On the model of the beryllium atom given by Bohr the configurations $4_{(3)}-4_{(2)}$ would appear to provide a spectral series with limits approximating to the above wave-lengths.

With boron only one critical voltage, namely 147·5, was observed in the higher range of applied potentials.

Hughes, too, found but one critical potential for boron in this range, namely 148 volts. The results obtained with lithium and beryllium would suggest that this failure to observe more than one critical potential could arise from boron's possessing a spectral series with a frequency formula $\nu = \mathrm{N}(1-1/m^2)$, with its first and last members not far apart. On this basis we may take 147·5 or 148 as approximating very closely to the quantum voltage of the shortest member of the series. The value of this wave-length would be approximately 83·6 A.

As regards carbon, a number of different critical potentials have been found by various investigators with the use of either the photo-electric method or the method of Frank and Hertz.* Kurth gives 289·4 volts, Richardson and Bazzoni 286 volts, Mohler and Foote two, namely, 272 volts and 234 volts. Finally, Hughes gives 215 volts. These, it will be seen, cover a range of wave-lengths extending from $\lambda = 57{\cdot}5$ A. to $\lambda = 42{\cdot}6$ A. The mean of all the values given above is 259·3 volts, and this by the quantum relation corresponds to $\lambda = 47{\cdot}6$ A. It would appear, therefore, that $\lambda = 47{\cdot}6$ A. must be close to the value of the wave-length of the shortest member of the K series for carbon.

The next element for which we appear to have any authoritative information is nitrogen. For this element Mohler and Foote found critical potentials of 352 and 374 volts. If we take the larger of these as representing the shortest member of the K series for nitrogen, it would give a limiting wave-length for the series approximating to $\lambda = 33$ A.

---

* Frank and Hertz, 'Ber. d. Deutsch. Phys. Ges.,' p. 457 (1914); Bergen Davis and Goucher, 'Phys. Rev.,' vol. 10, No. 2 (1917); Mohler and Foote, *loc. cit.*

For oxygen, Kurth gives the last member of the K series as $\lambda = 23{\cdot}8$ A., with a corresponding critical potential of 518·1 volts. For the same element Mohler and Foote found a critical potential of 478 volts, corresponding closely to $\lambda = 25{\cdot}8$ A.

These two results, which appear to be the only ones observed, give a mean critical potential of 498 volts. This value would give $\lambda = 24{\cdot}76$ A. as an approximation for the last member of the K series for oxygen.

Table VII.—Critical K Series Absorption Wave-lengths.

Element.	At. No.	Critical potential V. (volts).	$\sqrt{V}$.	Critical absorption wave-length × 10^8.
Potassium	19	3590·3	59·9	3·44 cm.
Argon	18	3189·8	56·5	3·86
Chlorine	17	2812·5	53·0	4·38
Sulphur	16	2460·1	49·6	5·01
Phosphorus	15	2141·5	46·3	5·76
Aluminium	13	1552·0	39·4	7·95
Magnesium.........	12	1296·0	36·0	9·51
Oxygen	8	498·0	22·3	24·76
Nitrogen	7	374·0	19·34	33·0
Carbon	6	259·3	16·1	47·6
Boron	5	147·5	12·15	83·6
Beryllium	4	104·3	10·2	118·2
Lithium	3	42·4	6·5	290·8
Helium	2	25·4	5·05	485·5
Hydrogen	1	13·5	3·67	913·4

For fluorine, neon, and sodium there appear to be no experimental data available to assist us in ascertaining the limits of the K series for these elements. For magnesium and the elements of higher atomic weight evidence is available in the results of direct determinations of their critical absorption K series wave-lengths. These results are all given in a convenient form by Duane,* and a limited number of them are recorded in Table VII of this paper. The Table also contains the values we have found for the limiting wave-lengths of the K series of lithium, beryllium, and boron, and values provided by the results of Richardson and Bazzoni, Kurth, Hughes, or Mohler and Foote for carbon, nitrogen, and oxygen. The Table also contains the atomic numbers of the respective elements and the square roots of the critical voltages corresponding to the critical absorption wave-lengths. These, it has been shown by Duane, may be taken as approximately close to the critical emission K series wave-lengths as well. The ionisation potentials for helium and hydrogen, with their corresponding quantum wave-lengths, are also given in the Table.

---

* Duane, 'Bulletin of the Nat. Res. Council, U.S.A.' vol. 1, Part VI, No. 6, p. 383.

The curve shown in fig. 10 was plotted with the atomic numbers of the elements taken as abscissæ and the square roots of the critical voltages given in Table VII as ordinates. From the form of Curve 1 it is evident that for the elements cited from potassium to beryllium the square roots of the critical potentials are very closely proportional to the respective atomic numbers. For lithium, helium, and hydrogen the square roots of the critical voltages show a departure from the linear relation. These departures are, however, gradual and regular from beryllium to hydrogen, which goes to confirm the view that it is now customary to take regarding the spectra of hydrogen, namely, that the Lyman series is the K series for this element.

The numbers given in Table VII and plotted in fig. 10 enable one to calculate approximately the critical voltages for the elements for which the critical absorption K series wave-lengths (and also approximately the critical emission K series wave-lengths) have not as yet been found experimentally. These calculated values are given in Table VIII.

Table VIII.

Element.	Atom. No.	Calc. critical potential (volts).	$\sqrt{V}$.	Calc. critical absorption or emission wave-lengths.
Silicon...............	14	1831·8	42·8	$6{\cdot}73 \times 10^{-8}$ cm.
Sodium	11	1075·8	32·6	11·46
Neon	10	852·6	29·2	14·46
Fluorine	9	670·8	25·9	18·4

VI. Low Range Critical Potentials and L Series Spectra.

(a) *Lithium.*

In studying the radiations emitted by lithium when bombarded by electrons it was found that in the lower range of applied potentials a marked discontinuity was always obtained for 12·0 volts. With lithium the radiation from the bombarded metal was not very strong when the accelerating fields were low, and on this account one could not be very sure whether any discontinuities existed or not below 12·0 volts. From his model of the atom of this element Bohr has obtained an energy change of 0·89 W_0 for the change of state $3_{(2,1)}-3_{(2)}$. If W_0 be taken equal to 13·5 volts, the ionisation potential of hydrogen atoms, we find 0·89 W_0 is equal to 12·02 volts. The agreement between this result and the value of the critical potential noted by us, namely 12·0 volts, suggests that the discontinuity we observed originated in the emission by lithium of a radiation of the L series type.

Since the change of state represented by $3_{(2,1)}-3_{(2)}$ involves the removal of the outer ring electron from the atom of lithium, it follows that the energy change corresponding to it, *i.e.*, 12·0 volts, should be the quantum equivalent of the wave-length of the shortest member of the L series. If we follow the general practice and take the L series to be given by a frequency formula of the form $\nu = \mathrm{N}(1/2^2-1/m^2)$, the quantum equivalent of the first member of the series would be 5/9ths of 12·0 volts or 6·7 volts. With this interpretation of the discontinuity noted by us, it would follow that the L series for lithium, should begin at about $\lambda = 1840$ A., and extend approximately to $\lambda = 1019$ A. Such a series has not as yet been observed spectrographically for lithium, but with a fluorite or vacuum grating spectrograph it should not be difficult to photograph it if it be really emitted by radiating atoms of lithium.

(b) *Beryllium.*

With beryllium discontinuities were obtained with the lower range of applied potentials at 16 and 20·3 volts. As these voltages are approximately in the ratio 20 : 27, this suggests that these may be interpreted as representing the first two members of the L series for beryllium provided such a series has a frequency formula of the form $\nu = \mathrm{N}(1/2^2-1/m^2)$. On the assumption that this is the correct interpretation of the discontinuities mentioned it follows that the calculated value of the quantum equivalent of the last member of the series is 28·8 volts. The wave-lengths of the L series for the element should begin therefore at about $\lambda = 770·7$ A., and extend approximately to $\lambda = 428·1$ A.

From his model of the beryllium atom Bohr has shown that the energy change involved in the change of state $4_{(2,\,2)}-4_{(2,\,1)}$ is equal to 1·95 W_0. If we take, as before, W_0 equal to 13·5 volts, the energy change referred to would be equivalent to 26·3 volts. The closeness of this value to 28·8 volts would furnish a warrant therefore for concluding that the discontinuities we found at 16 and 20·3 volts arose from the consecutive emission by the bombarded beryllium of the first two members of its L series spectrum.

(c) *Boron.*

With the lower range of applied potentials in the case of boron discontinuities were observed at 23·45 and 27·92 volts. These voltages also are approximately in the ratio 20 : 27, and if we follow the procedure adopted in the case of beryllium, we may take them to be the quantum equivalents of the first two members of the L series for boron. On this assumption the calculated quantum equivalent of the last member of the series would then be 42·2 volts, and the wave-lengths of the series should begin at about $\lambda = 525·8$ A. and end approximately at $\lambda = 292·2$ A. As Bohr in his paper did not discuss in detail his model of the atoms of boron, no results are available with which to check his conclusions. Considerations will be presented later in this paper, however, which show that our results for this element fit in well with those we obtained for lithium and beryllium, and also with data derived from investigations by others on the radiations emitted by elements of still higher atomic weight.

(d) *Carbon.*

It will be recalled that Hughes in his investigation of carbon found a discontinuity in his photoelectric curve at 34·5 volts. In discussing this result he presented a number of reasons, based on considerable collateral evidence, which led him to conclude that this voltage was the quantum equivalent of the critical absorption L series X-ray wave-lengths of carbon. By the quantum relation this would correspond to $\lambda = 358$ A. Kurth, too, found a discontinuity in his photo-electric curves at 32·9 volts, and concluded that $\lambda = 375$ was the convergence wave-length in the L series of this element. Mohler and Foote, however, give 75 volts as the lowest critical voltage observed by them below 234 volts. This, it will be noted, is the quantum equivalent of approximately $\lambda = 165$ A. In some preliminary observations that we have made distinct discontinuities were observed in our curves at about 33 and 72 volts.

Millikan* in his experiments on photographing the spectra of the elements in the extreme ultra-violet region found that the spectrum of carbon terminated abruptly at $\lambda = 360·5$ A. With other elements he found spectra extending considerably below this limit, which showed that his grating was able to register waves shorter than $\lambda = 360·5$ A., if they were emitted by carbon, even with moderate intensities. If, in view of the results obtained by Hughes, Kurth and Millikan, a wave-length approximately equal to 360 A. be taken as the short wave limit of the L series of carbon, the question of the significance of the wave-length $\lambda = 165·0$ A. connoted by the discontinuity at or about 75 volts arises. From considerations already presented it would appear that the convergence wave-length of the K series of carbon is approximately 47·6 A. If the frequency formula for the K series of carbon is of the form $\nu = N(1-1/m^2)$, a form that has been found by Moseley to apply for the heavy elements, it would follow that the wave-length of the first member of the K series for carbon should be approximately 63·5 A., with a quantum equivalent of 194·2 volts. The wave-length $\lambda = 165·0$ A. cannot, therefore, on this view belong to the K series for carbon. Another view that may be taken is that a wave-length at or near $\lambda = 165·0$ A. may be taken to be approximately the convergence wave-length of the L series for carbon, and that the discontinuities noted by Kurth and Hughes and ourselves between 33 and 34·5 volts may connote the first rather than the last member of the L series. This view, however, would be directly opposed to the conclusions of Millikan, who strongly emphasized the point that his spectrograms for carbon show an abrupt termination of wave-lengths at 360·5 A.

---

* Millikan, 'Ast. Phys. Jl.,' vol. 52, No. 1, p. 47 (July, 1920).

In considering this point, however, one is reminded that in photographing series spectra generally (including X-ray series) one rarely, if ever, is able to obtain photographic records of the members of a series corresponding to the higher frequencies. It is practically always the case that the first and following low-frequency members of the series possess the stronger intensities. It is difficult from evidence as yet available to settle this point definitely, but we venture to put forward the view that the discontinuity Hughes and Kurth and we observed in the photo-electric current at or near 33–34·5 volts was possibly the effect of the emission by carbon of the first member of its L series, and that the discontinuity noted by Mohler and Foote and by ourselves at or near to 74 volts (mean value) represented the emission of carbon of a radiation corresponding approximately to the convergence wave-length of the L series of the element.

(e) *Oxygen.*

In regard to the element oxygen, Kurth appears to be the only investigator who has observed a critical potential that can be connected with the L series of the element. His photo-electric current curves show a discontinuity at 49·72 volts, with a quantum equivalent of $\lambda = 248$ A. This wave-length he has taken to be the convergence wave-length of the L series for oxygen.

(f) *Aluminium and Silicon.*

With aluminium and silicon Kurth found discontinuities at 123·3 volts and 149·5 volts respectively, and the corresponding wave-lengths $\lambda = 100{\cdot}0$ A. and $\lambda = 82{\cdot}5$ A. he took to be the converging values of the L series wave-lengths for the two elements.

(g) *Phosphorus.*

With phosphorus, critical potentials were observed by Mohler and Foote at 95 volts ($\lambda = 130$ A.), 110 volts ($\lambda = 112$ A.), 126 volts ($\lambda = 98$ A.), and 163 volts ($\lambda = 77{\cdot}1$ A.). In interpreting these critical potentials, Mohler and Foote suggest that 126 volts and 95 volts may be taken to connote L (a_1) and L (a_5) respectively. For the remaining ones no definite interpretation is offered.

(h) *Chlorine.*

For chlorine, Mohler and Foote observed discontinuities in the photo-electric current curves at 198 volts (λ = 62·3 A.) and 157 volts (λ = 78·6 A.). These wave-lengths they took to be L (α_1) and L (α_6).

In interpreting the critical voltages cited for oxygen, aluminium, silicon, phosphorus, and chlorine, we have taken the view that the voltages which the authors named have taken to correspond to the convergence wave-lengths of the L series for these elements, should be considered as representing the first rather than the last members of such series. All of the critical potentials

Table IX.—Critical Voltages and L Series Wave-lengths.

Element.	Atom. No.	Crit. potential V.	$\sqrt{V}$.	Wave-length × 10^8.	Crit. potential V.	$\sqrt{V}$.	Wave-length × 10^8.
		volts		cm.	volts		c.m.
Lithium	3	6·7	2·6	1840·0	12·0	3·47	1019·0
Beryllium ...	4	16·0	4·0	770·7	28·8	5·36	428·1
Boron	5	23·45	4·84	525·8	42·2	6·5	292·2
Carbon	6	33·00	5·74	373·7	74·0	8·66	166·7
Oxygen	8	49·72	7·1	248·0			
Aluminium	13	123·3	11·1	100·0			
Silicon	14	149·5	12·23	82·5			
Phosphorus	15	163·0	12·8	75·6			
Chlorine ...	17	198·0	14·07	62·3			

found as yet in the lower range of applied voltages for lithium, beryllium, boron, carbon, oxygen, aluminium, and silicon are given in Table IX. In addition, one critical potential observed with phosphorus and one with chlorine is included. The Table also contains the atomic numbers of the elements, the square roots of the critical potentials, and the quantum equivalent wave-lengths of the latter.

The results are plotted in fig. 11. It will be seen the values of $\sqrt{V}$ fall into two sets of values, the one represented by Curve 3, and the other by Curve 2.

On the basis of the interpretation of our results that we have put forward above, Curve 2 would correspond to the convergence frequencies of the L series of the elements named, and Curve 3 to the frequencies of the first members of these series.

With the object of showing how this interpretation of the observed critical potentials fits in with the known values of the emission L series wave-lengths of the heavier elements, a limited number of these has been selected and they are given in Table X. The values given are those of the longest and of the shortest known L waves of the elements selected, and have been taken from Duane's Tables. The square roots of the quantum voltages of these waves are plotted against atomic numbers in fig. 12. The values given in Table IX are also plotted in the same figure, and, in order to make the record as complete as possible, the square roots of the additional critical voltages observed by Mohler and Foote for chlorine, sulphur, phosphorus, magnesium, sodium, and potassium, and by Richardson and Bazzoni for molybdenum, are marked on the diagram. These additional critical voltages are given in Table XI.

Table X.—L Series Wave-lengths.

Element.	Atom. No.	Longest observed wave-length.			Shortest observed wave-length.		
		$\lambda \times 10^8$.	V.	$\sqrt{V}$.	$\lambda \times 10^8$.	V.	$\sqrt{V}$.
		cm.	volts		cm.	volts	
Zinc	30	12·346 α_1	1000·0	31·6	—	—	—
Bromine	35	8·391 α_1	1469·5	38·3	8·141 β_1	1514·6	38·9
Rubidium ...	37	7·335 α_1	1681·1	41·0	7·091 β_1	1739·2	41·7
Strontium ...	38	6·879 α_1	1792·6	42·3	6·639 β_1	1857·4	43·1
Zirconium ...	40	6·083 α_1	2027·1	45·0	5·386 γ_1	2280·4	47·9
Molybdenum	42	5·410 α_2	2279·3	47·7	5·175 β_1	2382·8	48·8
Palladium ...	46	4·374 α_2	2819·1	53·1	3·597 $\gamma_2-\gamma_3$	3150·1	56·1
Cadmium ...	48	3·959 α_2	3114·7	55·8	3·328 γ_1	3705·2	60·9
Tin	50	3·604 α_2	3421·5	58·5	2·831 γ_4	4355·7	66·0
Antimony ...	51	3·443 α_2	3581·5	59·8	2·782 $\gamma_2-\gamma_3$	4432·4	66·6
Caesium	55	2·899 α_2	4253·5	65·2	2·234 $\gamma_2-\gamma_3$	5519·7	74·3
Cerium	58	2·578 α_2	4788·6	69·2	2·003 $\gamma_2-\gamma_3$	6156·25	78·5
Neodymium...	60	2·379 α_2	5183·3	72·0	1·775 γ_3	6947·04	83·4
Europium ...	63	2·131 α_2	5786·5	76·1	1·590 γ_3	7755·3	88·6
Erbium	68	1·794 α_2	6878·5	82·9	1·316 γ_3	9370·0	96·8
Tungsten	74	1·675 ϵ	7361·8	85·8	1·026 γ_4	12018·5	109·6

Table XI.—Miscellaneous Critical Potentials.

Element.	Atom. No.	V (volts).	$\sqrt{V}$.	$\lambda \times 10^8$.	Investigator.
Sodium	11	35	5·91	353	Mohler & Foote.
		17	4·12	726	
Magnesium ...	12	46	6·78	268	" "
		33	5·74	374	
Phosphorus ...	15	126	11·22	98	" "
		110	10·49	112	
		95	9·8	130	
Sulphur.........	16	152	12·33	81·2	" "
		122	11·05	101·0	
Chlorine	17	157	12·53	78·6	" "
Potassium......	19	23	4·8	537	" "
		19	4·36	650	
Molybdenum	42	356	18·9	34·8	Richardson & Bazzoni.

The diagram in fig. 12 brings out the well-known fact that for the heavier elements both the long L waves and the shorter ones as well follow closely the Moseley law $\sqrt{V} \propto$ At. No. The diagram also shows that for the

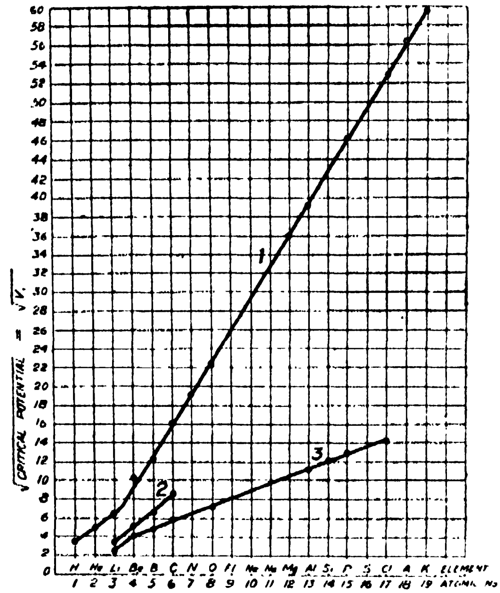

√CRITICAL POTENTIAL = √V.
0
2
4
6
8
10
12
14
16
18
20
22
24
26
28
30
32
34
36
38
40
42
44
46
48
50
52
54
56
58
60
1
2
3
H He Li Be B C N O Fl Ne Na Mg Al Si P S Cl A K ELEMENT
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 ATOMIC No

FIG. 11.

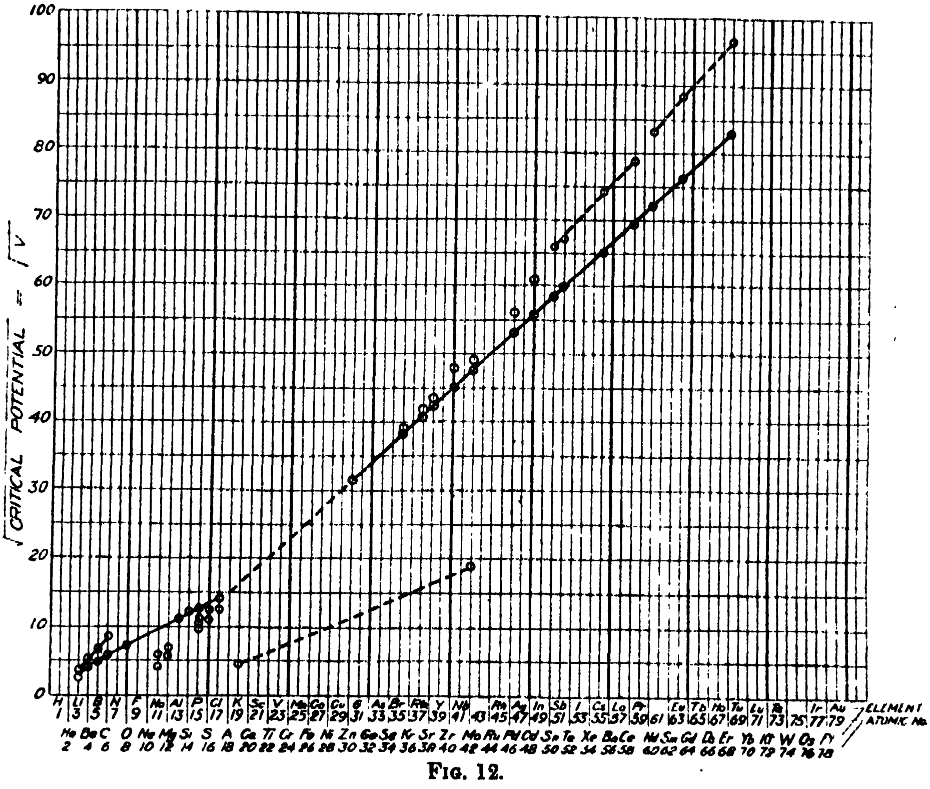

√CRITICAL POTENTIAL = √V
0
10
20
30
40
50
60
70
80
90
100
H Li B N F Na Al P Cl K Sc V Mn Co Cu G As Br Rb Y Nb Rh Ag In Sb I Cs La Pr Eu Tb Ho Tu Lu Ta Ir Au ELEMENT
1 3 5 7 9 11 13 15 17 19 21 23 25 27 29 31 33 35 37 39 41 43 45 47 49 51 53 55 57 59 61 63 65 67 69 71 73 75 77 79 ATOMIC No.
He Be C O Ne Mg Si S A Ca Ti Cr Fe Ni Zn Ge Se Kr Sr Zr Mo Ru Pd Cd Sn Te Xe Ba Ce Nd Sm Gd Ds Er Yb Kt W Os Pt
2 4 6 8 10 12 14 16 18 20 22 24 26 28 30 32 34 36 38 40 42 44 46 48 50 52 54 56 58 60 62 64 66 68 70 72 74 76 78

FIG. 12.

lighter elements from lithium to chlorine a similar law is followed. The proportionality factor between $\sqrt{V}$ and At. No., however, is not the same for the lighter elements as for the heavier ones. It is evident that a break occurs at or near chlorine or argon, but it is not clear why this should occur.

The electronic configuration suggested for argon by Bohr, based on the periodicity of the elements, consists of two electrons in the inner ring, eight in the second ring, and eight in the third ring. As the L rays have been supposed to originate in disturbances given to the electrons in the second ring, it would appear that when the third ring has its full quota of eight electrons it exerts a certain definite influence on the stability of the second ring of eight electrons. For elements whose third ring does not contain the full quota the stability of the second ring electrons is somewhat enhanced.

On the other hand it will be recalled that Moseley's empirical formula for the frequencies of the L-rays is of the form

$$V = 3{\cdot}29 \times 10^{15} (N - S_n)^2 (1 - 1/m^2) \text{ where } S_n = 7{\cdot}4.$$

According to Bohr's interpretation of this formula the value of S_n (namely 7·4) could arise from a ring of sixteen electrons surrounding the atomic nucleus. If for the heavier atoms the second ring did contain sixteen electrons, one would expect a modification of the Moseley law for the series spectra to appear similar to that indicated in fig. 12 when the element argon was reached. From Bohr's investigation it seems, however, that a ring of sixteen electrons would not be a stable configuration for atoms whose atomic number was less than 72.

With the experimental data at present available it does not seem possible to reach more definite conclusions as to the frequency limits of the L series spectra of the lighter elements, and the validity of the view that has been put forward in this paper can only be fully tested when the short wave-length radiations emitted by the lighter atoms have been more closely examined. From what has been put forward it seems clear that for the lighter elements from lithium to carbon certain critical voltages have been found for which the square roots are connected with the corresponding atomic numbers by two linear relations. In the one case the critical voltages constitute a group that appear to connote the convergence frequencies of the L series spectra of the elements, while in the other they form a group that we have taken to represent the frequencies of the first members of these series. It seems evident that the critical voltages of the second group fit in (according to the linear relation connecting them) with certain observed critical potentials for the elements oxygen, aluminium, silicon, phosphorus and chlorine. For elements heavier than chlorine, it would appear from fig. 12 this linear relation no longer holds and that Moseley's law then becomes operative. It is interesting to note that the linear relation connecting the second group of potentials mentioned merges into the Moseley law when the element argon is reached.

(i) *Critical Potentials and M Series Spectra.*

In discussing their results, Mohler and Foote have taken the view that the critical potentials observed by them with potassium at 19 volts and 23 volts should be identified with the M series spectrum of this element.

Richardson and Bazzoni have also expressed the opinion that there is good ground for identifying the critical potential observed by them with molybdenum at 356 volts with the M X-rays of this element.

Even if these views should turn out to be correct, it still leaves a number of the critical potentials given in Table XI unidentified. From the representation of them shown in fig. 12 a number of them would appear to be connected by linear relations, but it does not seem possible at present to reach definite conclusions in regard to their significance.

(j) *Summary of Results.*

1. Critical potentials have been determined for the elements lithium, beryllium, and boron which correspond to the K series critical absorption X-ray wave-lengths of these elements.

2. It has been shown that the Moseley linear law connecting the square roots of the frequencies of the critical absorption K series wave-lengths with the atomic numbers of the elements applies not only for the heavier elements, but also for the lighter ones down to lithium.

3. The results obtained fit in with the view that the Lyman ultra-violet series for hydrogen is the K X-ray series of this element, and with the view that the convergence wave-length of the K series for helium is approximately $\lambda = 485{\cdot}5$ A.

4. Additional series of critical potentials have been determined for lithium, beryllium, boron, and carbon. One set of these potentials has been interpreted as representing the convergence wave-lengths of the L series for these elements, and another set as representing the first members of the L series for the same elements. In both cases the square roots of the critical voltages have been found to be directly proportional to the atomic numbers of the elements.

5. Evidence has been adduced in support of the view that the Moseley law, which is known to apply for the L series of the heavier elements, ceases to apply for elements lighter than argon, and that the linear relation connecting

the atomic numbers of these lighter elements with the exciting voltages of their L series merges into the ordinary Moseley law when the element argon is reached.

In conclusion, we wish to take this opportunity of acknowledging our indebtedness to Mr. J. F. T. Young for kindly assisting us at various times during the investigation.

[*Note added November* 15, 1922.—The attention of the writers was recently drawn to a paper by Prof. Millikan,* describing the results of a series of brilliant experiments in which grating spectra of the light elements were photographed in the extreme ultra-violet as far down as λ 136·6 A. In this paper Prof. Millikan draws the conclusion that the Moseley law for L series spectra applies very closely for all elements from uranium to neon, and that for the elements lighter than neon, the square roots of the voltages corresponding to the L series critical absorption wave-lengths lie on a line which is slightly concave towards the atomic number axis, but which meets the line representing Moseley's law for the heavy elements at the element neon.

Acceptance of this view would mean that the line in fig. 12, passing through the points representing the critical potentials for the elements from zinc to erbium, should be extended so as to pass through the group of unconnected points shown in the diagram, as representing certain experimentally determined critical potentials (see Table XI) for the elements sodium, magnesium, phosphorus, sulphur and chlorine. If this view should prevail, the critical potentials found by the writers and discussed in the present paper, as well as others found by Richardson and Bazzoni, Hughes, Kurth, Mohler and Foote, would be left unexplained.

On the other hand, the correctness of Prof. Millikan's interpretation of his photographically recorded spectra of the light elements would be called in question if the interpretation we have ventured to place upon our results be accepted.

I desire to state that Miss M. L. Clark was enabled to co-operate in carrying out this investigation through the award to her of a bursary kindly made by the Honorary Advisory Council for Scientific and Industrial Research of Canada.—J. C. M.]

---

* R. A. Millikan, 'Proc. Nat. Acad. of Sc. of the U.S.A.,' vol. 7, No. 10, p. 289 (1921).

A Theory of Meteors, and the Density and Temperature of the Outer Atmosphere to which it Leads.

By F. A. LINDEMANN, F.R.S., Professor of Experimental Philosophy, and G. M. B. DOBSON, M.A., Lecturer in Meteorology, University of Oxford.

(Received October 21, 1922.)

INTRODUCTION.

Some thirty years ago Teisserenc de Bort announced the discovery that the temperature gradient in the atmosphere, which amounts to some $6{\cdot}5 . 10^{-5}$ degree/cm. for the first 10 km., becomes negligibly small at greater heights. Since this time it has become usual to treat the temperature as constant for all heights above this limit, and various detailed numerical estimates, extending to heights of hundreds of kilometres, have been published, which are based on this assumption. So far, however, no evidence for it has been adduced beyond the "ballon sonde" observations, which scarcely extend above 25 km.

The purpose of this note is to examine how much information about conditions at greater heights may be gained from a study of meteors. For this purpose it is necessary to study in considerable detail the theory of what happens when a meteor appears. It will be seen that a consistent theory can be developed, which accounts for all the observed phenomena, and that this theory enables one to derive a considerable amount of information about conditions in the upper air. It will be shown that existing observations enable us to say with considerable certainty that the density at heights above 65 kilom. is very much higher than is commonly supposed, and that the temperature must increase from its value of something like 220° abs. at heights between 12 and 50 km., to something like 300° abs. at these heights. It is evident that accurate results can only be obtained when the required quantities, viz., heights of appearance and disappearance, length of path, velocity and luminosity have been determined for a number of meteors. It is hoped that experiments, now in progress to determine these quantities photographically, will enable more accurate calculations to be made. In the meantime, the broad results obtained from the data due to W. F. Denning and his co-workers may be outlined.*

---

* For Table of Symbols, see p. 436.

Part I.—METEORS.

The heights, paths and velocities of some thousands of meteors have been observed. They appear, in general, at heights between $1{\cdot}6 \,.\, 10^7$ and $7 \,.\, 10^6$ cm., and disappear at heights below $1{\cdot}2 \,.\, 10^7$ cm., mostly at about $8 \,.\, 10^6$, though some, of course (the so-called meteorites) reach the ground. Velocities from $9{\cdot}5 \,.\, 10^5$ to $1{\cdot}6 \,.\, 10^7$ cm./sec. have been recorded, though these latter are probably not accurately observed, since they could only occur if the particles were not components of the solar system. The path may be at any angle to the horizontal, from almost horizontal to vertically downward. Enormous variations in brightness occur, from a brightness well beyond that of the full moon down to a brightness estimated as equal to that of a seventh magnitude star, in telescopically observed meteors. For the present we shall exclude fire-balls and the like, and confine our attention to the ordinary meteors. For convenience in checking the order of magnitude to which formulæ lead, we will typify these by a shooting star appearing at a height of 10^7 cm. and disappearing at a height of $8 \,.\, 10^6$ cm., after traversing a path of $6 \,.\, 10^6$ cm., at a speed of $4 \,.\, 10^6$ cm./sec. We will assume further that its average apparent brightness at a distance of $1{\cdot}5 \,.\, 10^7$ cm. is equal to that of a first-magnitude star, *i.e.*, that the energy $E = 3{\cdot}3 \,.\, 10^{10}$ ergs is radiated per second.

It will appear that the only possible explanation of these phenomena is that usually accepted, namely, that the meteor appears when a cosmic particle of matter, moving at a sufficiently high speed relative to the earth, becomes heated by atmospheric friction to such a temperature that it evaporates, and that it disappears when all the matter originally present has evaporated. The molecules which distil off the meteor, as also the molecules of air in its path after collision, are moving at approximately the speed of the meteor, and lose their kinetic energy, largely in the form of radiation, by collision with other atmospheric molecules.

Head-resistance of Particle.

All meteors move at a speed great compared to the velocity of sound, and it is therefore essential to examine the question of the amount of head-resistance encountered by a particle moving through a gas at such a velocity. Strictly speaking, of course this must depend upon the shape of the particle, since the head-resistance must be diminished if the air can easily flow round the particle. At high speeds, however, only that part of the air which strikes the particle at an angle less than V/v can flow away, V being the velocity of sound and v that of the particle. Part of the air, which impinges at a slightly greater angle, will not reach the full velocity, but practically all the air in the path of the meteor will be accelerated to the meteor's velocity. As long as v is great compared to V, therefore, one may put the amount of air accelerated equal to $\mathrm{S}\int \rho_0 v\,dt = \mathrm{S}\int \rho_0\,dl$, S being the cross-section and ρ_0 the density.

The total kinetic energy abstracted from the meteor per second, *i.e.*, the work done against the head-resistance of the air, is converted either into kinetic or potential energy of the air, the latter of which appears in the form of adiabatic compression and heating. The kinetic energy is clearly $m'v^2/2 = \mathrm{S}\tfrac{1}{2}\rho_0 v^3$, since $m' = \mathrm{S}\rho_0 v$. The work done in compressing the air adiabatically from the pressure p_0 to the pressure p is $m'(\mathrm{RT}_0/\mathrm{M}_0)\log p/p_0$, M_0 being the molecular weight. As $\mathrm{RT}_0 = \tfrac{1}{3}\mathrm{M}_0\mathrm{V}_0^2 = \mathrm{M}_0\mathrm{V}^2/\gamma$, γ being the ratio of the specific heats of the gas, the potential energy is therefore

$$(m'\mathrm{V}^2/\gamma)\log p/p_0 = (\mathrm{S}\rho_0 v\mathrm{V}^2/\gamma)\log p/p_0.$$

The pressure p is given in terms of the work per second E by E/Sv. Since E is the sum of the kinetic and potential energy, therefore

$$p = \tfrac{1}{2}(\rho_0 v^2)\{1 + (2/\gamma)(\mathrm{V}^2/v^2)\log p/p_0\}.$$

It is clear that the second term can be neglected when the velocity v is great, compared to the velocity of sound V, as in the case in meteors. Hence one can put, with sufficient accuracy, $p = \tfrac{1}{2}(\rho_0 v^2)$, and the head-resistance $\mathrm{F} = \mathrm{S}\,\tfrac{1}{2}(\rho_0 v^2)$.

Temperature of Surface of Particle.

A point of considerable importance may readily be settled once this formula is known, namely, the temperature to which the surface of the particle is heated by the friction of the gas. It is clear that the layer of gas in front of the meteor will be heated by adiabatic compression. The surface temperature of the particle will approach the temperature of the compressed gas which, of course, it cannot exceed. The temperature of the gas is given by

$$\mathrm{T}_1/\mathrm{T}_0 = (p/p_0)^{(\gamma-1)/\gamma} = (3v^2/2\mathrm{V}_0^2)^{(\gamma-1)/\gamma} = (\gamma v^2/2\mathrm{V}^2)^{(\gamma-1)/\gamma} = (\mathrm{M}_0 v^2/2\mathrm{RT}_0)^{(\gamma-1)/\gamma},$$

T_0, of course, being the initial temperature of the gas and M_0 its molecular weight. Thus, for instance, for N_2, if $\mathrm{T}_0 = 220°$ and $v = 4\,.\,10^6$, $\mathrm{T}_1 \underset{.}{=} 3420°$. For $v = 10^6$, however, T_1 would barely exceed 1550°. The temperature of the surface will, of course, be somewhat lower than the temperature of the gas. If the rate of evaporation is known this difference may be found by equating the latent heat, *i.e.*, the rate at which heat is conducted into the particle, to the heat-flow from the gas for a given difference of temperature.

Flow of Heat into Particle from Gas.

The rate at which heat flows from the compressed gas into the particle must be considered for two main cases. In the first which occurs at low densities, the gas molecules in general impinge directly upon the surface of the particle, and do not spend any of their energy in collisions with other gas or vapour molecules. In the second case at higher densities the free path is small and the ordinary gas laws hold.

The first case is not easy to treat, but it is possible to determine approximately within which limits it obtains. At the high speeds with which we are concerned, the accommodation coefficient fortunately cannot intervene. The velocity of the molecules after collision must be of the same order as the velocity of the meteor, and consequently the work done per second must be $S\frac{1}{2}(\rho_0 v^3)$, since the mass $S\rho_0 v$ is accelerated to a velocity v.

The difficulty lies in determining in what form this energy appears. The collision of even a hydrogen atom moving at $4 . 10^6$ cm./sec. corresponds to a temperature of something like 130,000°, so that part of the energy, probably the major part, will be used in breaking up the molecule and ionising the atoms; part of the energy again will be immediately radiated; a fraction will probably flow into the relatively cold meteor and tend to heat it; lastly, part of the energy may cause local evaporation of meteoric molecules. It is extremely difficult to form any accurate idea of the relative importance of these four terms, but a rough consideration, treating the solid as continuous, and using Stefan's law, appears to indicate that only a minor portion of the energy will heat or evaporate the meteor. For the heat cannot spread faster than with the velocity of sound, and it may readily be shown that practically all the energy generated by collision with one atom would be radiated away before the sound-wave could travel as far as the neighbouring atoms. Though the method is admittedly inaccurate, since the ordinary laws of black body radiation do not hold for single molecules, it probably justifies one in assuming that the meteor is not heated appreciably during this first stage.

If this is so it is comparatively easy to determine the density at which the meteor will be protected by a layer of gas and direct impacts cease. This will occur when the average free path is equal to the distance covered by the meteor in the average time taken by a molecule to escape after collision from the region swept by the meteor, *i.e.*, when L_1 is of the order $r . v/V_0$, r being the radius and v the velocity of the meteor, V_0 that of the molecule. If L_0 is the free path at normal temperature and pressure, therefore direct impacts will occur up to a density ρ_1 which is $L_0/r . V_0/v$ times the standard density. For H_2 this reduces to $3.10^{-4}/rv$, for He to $6 \cdot 7.10^{-4}/rv$, and for N_2 to $5 \cdot 6.10^{-4}/rv$, if V_0 is taken as corresponding to a temperature of 300°, which is certainly a minimum value which can be assumed. Taking $5 \cdot 6.10^{-4}$ as H_2 and He occur only in small quantities at the normal height of appearance, therefore, the first case, that of direct impact, obtains up to densities of the order $5 \cdot 6.10^{-4}/rv = 1 \cdot 4.10^{-10}/r$ for the assumed average meteor. If some molecules are volatilised by the impacts of course they will increase the effective density, and hence form a protective layer at lower actual gas pressures than correspond to the above density. If many evaporate the meteor might become visible even in this region. As stated above, however, the fraction of the heat generated which becomes available for raising the temperature of the meteor is probably negligible.

The second case can be treated more accurately. Here the density is so great that the ordinary gas laws apply, the incoming molecules do not impinge directly upon the surface of the solid, but upon a couch or cap of compressed air and vapour built up in front of the meteor.

If the head-resistance of the particle is $F = S \frac{1}{2} (\rho_0 v^2)$, the work done per second in moving a distance v is, of course, $S \frac{1}{2} (\rho_0 v^3)$. Only a fraction of this work will, however, become effective in heating the meteor, since the greater part will be carried off and finally degraded into heat by the air and vapour leaving the meteor. Though the pressure on the front of the meteor is high, namely, $\frac{1}{2} (\rho_0 v^2)$, the velocity V' of the surface relative to the gas exerting this pressure is not high. Hence the heat actually flowing into the meteor, *i.e.*, the heat which becomes available for heating or evaporating the meteor, will be $\frac{1}{2} (\rho_0 v^2) V'$, and it is readily seen that V' cannot exceed the velocity of the gas molecules.

It is not difficult to derive an accurate expression for V' from the fundamental conceptions of the kinetic theory of gases. It is clear for this purpose, as indeed for all considerations in this paper, that it is permissible to neglect the law of the distribution of molecular velocities, and to take all molecules to be moving with a mean velocity $V_0 = \sqrt{(3RT_0/M_0)}$. The layer of gas immediately in contact with the meteor has been compressed from a pressure p_0 and temperature T_0 to a pressure $p = \rho_0 v^2/2$, and a temperature $T_1 = T_0(p/p_0)^{(\gamma-1)/\gamma} = T_0(3v^2/2V_0^2)^{(\gamma-1)/\gamma}$. If one calls the temperature of the meteor's surface T_2 the molecules will leave the surface with a velocity $V_2 = \sqrt{(3RT_2/M_0)}$. If one calls ν_1 the number of molecules in unit volume which have a component of velocity towards the meteor, the number striking unit surface in unit time will be $\nu_1 V_1 \int_0^{\frac{1}{2}\pi} \sin\phi \cos\phi d\phi = V_1 \nu_1/2$. This must clearly be equal to $V_2 \nu_2/2$ the number which leave the surface per second, since the number of vapour molecules leaving the surface is small compared to the number of collisions. A molecule arriving at an angle ϕ to the normal imparts the momentum $m_1V_1\cos\phi$ and leaving at an angle ψ, the momentum $m_1V_2\cos\psi$. Hence the total momentum transferred, *i.e.*, the pressure

$$p = m_1V_1 \,.\, V_1\nu_1\int_0^{\frac{1}{2}\pi}\sin\phi\cos^2\phi\,d\phi + m_1V_2 \,.\, V_2\nu_2\int_0^{\frac{1}{2}\pi}\sin\psi\cos^2\psi\,d\psi$$
$$= \tfrac{1}{3}m_1V_1\nu_1(V_1+V_2).$$

Since this is equal to $\frac{1}{2}\rho_0v^2$ therefore $V_1\nu_1/2$ the number of molecules striking unit surface in unit time is $\frac{1}{2}V_1\nu_1 = \frac{3}{4}n_0\,v^2/(V_1+V_2)$.

All that remains is to determine the mean velocity of all the molecules striking and leaving the surface, *i.e.*, the mean velocity of the gas relative to the surface. If the surface is inclined at an angle θ to the direction of motion the component of velocity at right angles to the surface is $V'\sin\theta$. Therefore the effectiveness of each element of surface $d\sigma$ is reduced to the fraction $\sin\theta d\sigma$ and the effective area of a solid where θ varies from 0 to $\pi/2$ is reduced from $\int d\sigma$ to $\int\sin\theta d\sigma$, *i.e.*, the effective surface is equal to the cross-section of the solid. For any molecule arriving at an angle ϕ there must be one leaving at an angle ϕ. For each such pair the velocity of the surface relative to the centre of gravity is $(V_1-V_2)/2 \,.\, \cos\phi$. The chance that any collision will take place at an angle between ϕ and $\phi+d\phi$ is $2\sin\phi\cos\phi d\phi$ so that the mean relative velocity $V' = (V_1-V_2)\int_0^{\frac{1}{2}\pi}\sin\phi\cos^2\phi d\phi = (V_1-V_2)/3$.

Hence the heat which flows into the meteor is

$$SpV' = S\,\tfrac{1}{9}(m_1V_1\nu_1)(V_1^2-V_2^2) = S\,\tfrac{1}{3}(m_1V_1\nu_1) \,.\, (R/M_0) \,.\, (T_1-T_2)$$
$$= S\,\tfrac{1}{2}(\rho_0v^3) \,.\, (V_1-V_2)/3v.$$

For many purposes it is convenient to introduce the idea of an efficiency factor for heating, *i.e.*, a coefficient giving the fraction of the total work done $S\frac{1}{2}(\rho_0v^3)$ which goes to heat the meteor. If we call this factor k therefore we may put

$$k = (V_1-V_2)/3v = \tfrac{1}{3}[\surd(T_1/T_v)-\surd(T_2/T_v)]. \qquad (1)$$

T_v being a fictitious temperature given by $M_0v^2/3R$. This factor is plotted in fig. 1 for various values of T_2 taking T_0 as 300°.

Theoretical Explanation of Meteors.

Having developed these formulæ which will be required we may consider what a meteor is and how the attendant phenomena are caused.

The only tenable explanation is that a meteor is an extra-terrestrial particle which enters the air at high speed, that it becomes visible (owing to collision of the fast vapour molecules with air molecules) when its surface becomes hot enough to evaporate appreciably, and that it disappears when it has evaporated practically completely. Deceleration is small, and radiation from the particle itself is negligible.

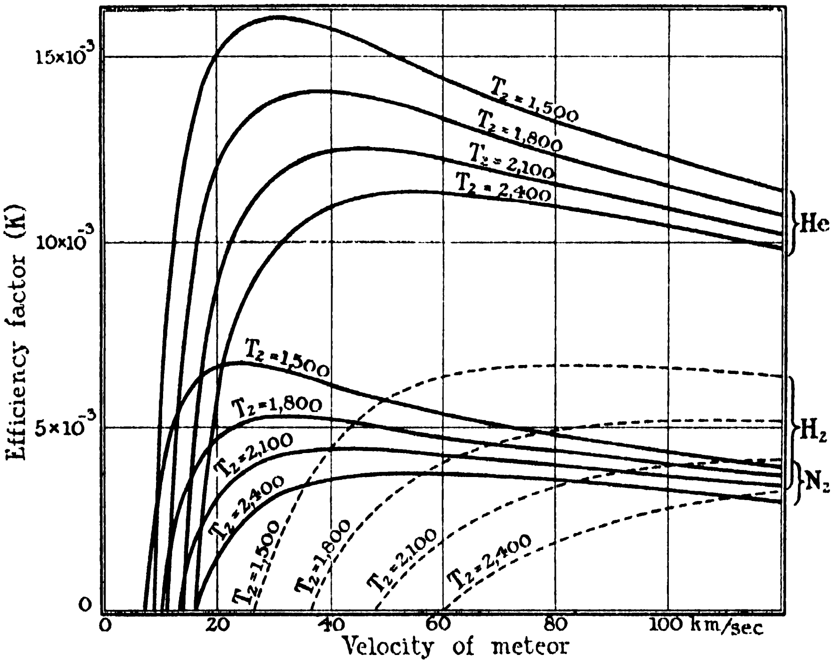


FIG. 1.

Only one possible modification of this view exists as far as we can see, and this is, that the meteor appears when the original cosmic particle liquefies and spray is formed by the air pressure. If the drops of spray are sufficiently minute they will evaporate as on the other theory and give rise to a luminous meteor. Their cross-section will, however, far exceed that of the original particle. Obviously this phenomenon really occurs in large meteors and meteorites, which reach comparatively dense strata. In these cases the actual fireball is followed by a luminous trail consisting of the spray blown away from the surface. The only question is at what densities and velocities this spray appears. In mercury a pressure of the order of 10^5 dynes/cm. is necessary, and it seems probable that a greater pressure would be required in the case of iron or olivine. In this case spraying cannot occur at great heights, except for very fast meteors. We are, therefore, justified in neglecting it in ordinary shooting stars, especially those which are not followed by luminous trails.

Form in which Kinetic Energy of Meteors Appears.

It appears certain that most of the energy of the particle is transformed into radiation. Collisions at speeds of the order of 10^6 cm./sec. are amply sufficient to ionise most atoms and even more, of course, to disturb the outer electrons. An iron atom moving at a speed of $3 . 10^5$ cm./sec. has the energy of a quantum of frequency $6{\cdot}3 . 10^{14}$ or of an ion after a fall through 2·6 volts. At these speeds, therefore, visible radiation will be produced by collisions. In the last stages of degradation the remaining fraction of energy may be carried away by ordinary heat conduction, but the major part of the energy corresponding to a drop in velocity from v to say $v/10$ will give rise to radiation. Another small fraction of the total energy will be used up in volatilising the meteor, but as this only corresponds, *e.g.*, in iron to a velocity of $3{\cdot}5 . 10^5$ cm./sec. it also is negligible in all cases.

Incidentally, this view is supported by the appearance of luminous trails, which remain visible sometimes for an hour after the meteor has passed. The energy required is so great that it can only be stored in electrostatic form, and there seems no alternative than to assume that these trails are caused by the slow recombination of ions.

A fraction of the heat produced returns to the particle in the form of radiation from the escaping gas, but this is negligible in most cases. The volatilised meteoric molecules, as also escaping air molecules, are moving with a velocity nearly equal to that of the meteor. As viewed from the meteor, they will appear as a shell expanding at a rate of the order V_3 and blown away by a wind of velocity v. They will, therefore, form a cone of solid angle, V_3/v. Neglecting absorption and integrating over the cone, the fraction of the total radiation which falls on the meteor is $\frac{3}{4} r^2/r_2^2 . V_1^2/v^2$. r_2 is defined by the fact that radiation is caused by the impacts of the air molecules, and that the escaping gas will be brought approximately to rest relative to the air and will have lost most of its energy when it has encountered a mass of air $\pi r_2^2 v\rho dt$ of the order of its own mass, dm. Since r_2 is of the order of 1 cm., as will be shown later, and v is large compared to V_1, $\frac{3}{4} r^2/r_2^2 . V_1^2/v^2$ is small in general, even compared to k.

Mass of Meteors.

Since it is possible to determine the total energy radiated, and practically all the energy appears in this form, it is easy to find the mass of the meteor if the velocity is known. Taking the figures assumed as average values above, *i.e.*, a meteor traversing a path of $6 . 10^6$ cm. in 1·5 seconds, and appearing at a distance of $1{\cdot}5 . 10^7$ cm. as bright as a first-magnitude star, $\int E dt = 5 . 10^{10}$ ergs and $v = 4 \,.\, 10^6$ cm./sec., whence $m = 2\int \mathrm{E}dt/v^2 = 6{\cdot}25 \,.\, 10^{-3}$ grm. If composed of iron, therefore, the diameter of such a meteor would be 0·115 cm. Much brighter meteors must be either correspondingly greater or faster. A meteor 10 magnitudes brighter, for instance, *i.e.*, of the order of the brightness of the moon, which moved at the same speed and lasted the same time, would have a mass of 62 grm. and a diameter of nearly 2·5 cm.

Deceleration of Meteoric Particles.

It is evident, as we have seen, that a particle moving through the air at a sufficient velocity will be heated and must consequently evaporate. It is easy to prove that such evaporation will under ordinary circumstances be almost complete before material deceleration occurs.

As was shown above, the heat which flows into the meteor is

$$\mathrm{S} \,.\, \tfrac{1}{2}\rho v^3(\mathrm{V}_1 - \mathrm{V}_2)/3v,$$

V_1 being the velocity of the compressed gas molecules in front of the meteor and V_2 that of the gas molecules at the temperature of the meteor's surface. Putting $\mathrm{S} = \pi r^2$ and $m = (4\pi/3) \,.\, r^3\rho_m$, ρ_m being the meteor's density, therefore

$$(\mathrm{V}_1 - \mathrm{V}_2) \,.\, \tfrac{1}{6}(\rho_0 v^2) \,.\, \pi r^2\, dt = l dm = 4\pi r^2 \rho_m\, l dr,$$

l being the latent heat per gramme, whence

$$dr = \tfrac{1}{24}(\mathrm{V}_1 - \mathrm{V}_2) \,.\, \rho_0/\rho_m \,.\, v^2/l \,.\, dt,$$

or since

$$dv = \mathrm{S}\,\rho_0 v^2/2m \,.\, dt = \tfrac{3}{8}\,\rho_0/\rho_m \,.\, v^2/r \,.\, dt,$$

$$(\mathrm{V}_1 - \mathrm{V}_2)\, dv = 9l\, dr/r,$$

or

$$v_0 - v_1 = 9l/(\mathrm{V}_1 - \mathrm{V}_2) \,.\, \log r_0/r_1,$$

if the variation of $\mathrm{V}_1 - \mathrm{V}_2$ with v be neglected.

For iron l is of the order $6 \,.\, 10^{10}$. Since V_1 and V_2 are of the order 10^5, it is evident that no material deceleration will take place until $\log r_0/r_1$ becomes important, say until r is less than $1/e$th of r_0, in other words, until 19/20 of the mass has evaporated. It will be seen later that in the slowest meteors V_1 is but little greater than V_2, and in these deceleration might be perceptible on the last part of the path, but for fast meteors it may be neglected.

Deceleration of Vapour.

Roughly speaking, this consideration has shown that a solid particle cannot be materially decelerated, since its velocity can only be reduced to $1/e$th when it has acted upon a column of air of twice its own mass, *i.e.*, when $\rho_0\mathrm{L}$ is of the order $\rho_m r$. Once the meteor evaporates, of course, it has an opportunity of interacting with a column of air of vastly greater cross-section. At each collision the velocity of the molecules will be reduced, so that the effective diameter will be of the order of 2 to 10 free paths, *i.e.*, at a density of 10^{-7} of the order of 0·2 to 1 cm. Another way to treat it is to calculate πr_1^2 by equating $\pi r_1^2 \rho_0 v$ to $2\,dm/dt$, for 0·86 of the energy will be lost by the time the cloud of vapour has expanded to this radius. In our typical case $r_1 = 2{\cdot}55 \, . \, 10^{-5}/\sqrt{\rho_0}$, in fair agreement with the estimate from the free path. Incidentally, the luminosity would extend to a distance v/V_1 times as great behind the meteor, but this would be too small to be perceptible. It may be of interest to mention here that one arrives at very similar value if one equates the total radiation to that of a black body at the optimum temperature and determines the radius of the radiating body from the equation $4\pi r_1^2 \sigma T^4 = E$.

Oxidation of Particles.

At first sight it might appear that meteors consisting of oxidisable material, *e.g.*, iron, might be appreciably heated by the energy of chemical combination. This could only occur in slow meteors at low heights. Fast meteors, as will be shown, will evaporate rapidly owing to friction, and atmospheric oxygen will scarcely penetrate to the surface through the escaping meteoric vapour in sufficient quantity to add materially to the surface heat. The vapour molecules will presumably eventually be oxidised, but this will only slightly affect the total energy emitted and not the duration, etc., of the meteor. If the temperature of the surface is comparatively low, however, as in the case in slow meteors, evaporation will be negligible, atmospheric oxygen will reach the surface, and the heat of chemical combination will be added to the heat flow due to friction. At sufficient densities this may become important, but it is evident that a number of oxygen molecules of the same order as the number of molecules in the meteor must be encountered before this occurs, since the latent heat of iron is of the same order per molecule as the heat developed by a molecule of oxygen combining with iron. As was shown above, the number of molecules which strike a surface πr^3 in unit time is $\pi r^2 \frac{3}{4} n v^3/(V_1 + V_2)$, n being the number in unit volume, v the velocity of the meteor, and V_1 and V_2 the velocity of incoming and departing molecules. If ρ_m is the meteor's density and M_1 the molecular weight, the number of molecules it contains will be $4\pi r_0^3 \rho_m/3 \, . \, N/M_1$, so that, leaving the change of r out of account, $\frac{3}{4} t \pi r_0^2 v^2 n/(V_1 + V_2)$ would have to be of the order $4\pi r_0^3 \, . \, N\rho_m/M_1$, *i.e.*, n of the order

$$16\pi^2(V_1 + V_2)/9v^4 \, . \, N\rho_m r/M_1.$$

For iron, $N\rho_m/M_1 = 8{\cdot}5 \, . \, 10^{22}$, so that for $t = 1{\cdot}5$, $v = 7 \, . \, 10^6$, $V_1 + V_2 = 1{\cdot}4 \, . \, 10^5$ and $r = 6 \, . \, 10^{-2}$, n would be of the order $1{\cdot}6 \, . \, 10^{16}$. This corresponds to an atmospheric density about 1/300 of the normal, such as obtains at heights, say, of $4 . 10^6$ cm. Above such heights oxidation would under no circumstances be important. At lower heights it might become apparent if the surface were not already evaporating owing to friction.

Evaporation of Particles.

It is evident that evaporation will take place if the temperature of the surface is raised to a sufficient value. This requires a certain speed, for since by a well-known equation, $T_1/T_0 = (p/p_0)^{(\gamma-1)/\gamma}$, one may write $T_1 = T_0(Mv^2/2RT_0)^{(\gamma-1)/\gamma}$. The subjoined diagram (fig. 2) gives the variation of T_1, with v for N_2, H_2 and He, T_0 being taken as 300 A. It is evident that, for meteoric velocities, T_1 will in general easily reach the required value.

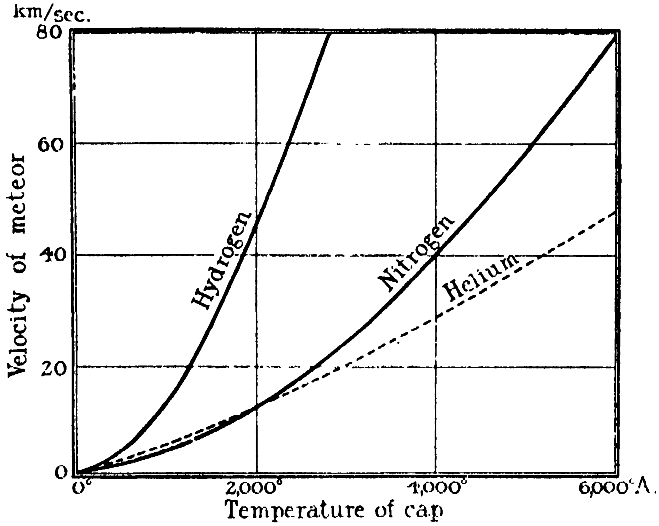


FIG. 2.

An estimate of the value of T_2, the temperature of the meteor's surface, can easily be made. The mass which evaporates into a vacuum from unit surface in unit time at vapour pressure p' must be equal to the mass which would strike unit surface in unit time under equilibrium conditions. Hence $dm/dt = n'm'V'/4 = \frac{3}{4}p'/V' = p'/4\sqrt{(3M'/RT)}$. Since the vapour is stopped rapidly, the energy radiated per second is $v^2/2 . dm/dt . 2\pi r^2$, *i.e.*, $E = \pi r^2 v^2 p'/4 . \sqrt{(3M'/RT')}$. In general, if $E_0 = \int E dt$, $m = 2E_0/v^2$, and $r = (3E_0/2\pi\rho_m v^2)^{1/3}$, as has been shown above. Taking a mean value for E such that $E_0 = Et_0$, t_0 being the meteor's duration, therefore

$$p' = (256/9\pi . E\rho_m^2/v^2 t_0^2)^{1/3} . (RT'/3M')^{1/2}.$$

E, in the case of our typical meteor, was $3{\cdot}3 . 10^{10}$, $v = 4 . 10^6$, $t_0 = 1{\cdot}5$.

Assuming it to be made of iron, therefore, $\rho_m = 7{\cdot}86$ and $M' = 56$, so that, putting T' of the order of 2000°, one finds $p' = 2{\cdot}52 \,.\, 10^4$, or about 19 mm. of mercury. It is clear that E may vary considerably, as also t_0, but, excluding extreme cases, it seems unlikely that p' should be more than 80 mm. or less than 4 mm.

If the meteor is composed of iron, these values of the vapour pressure enable one to calculate the temperature with the required accuracy. Taking the most probable vapour pressure curve, it seems likely that T_2 must lie between 2000° and 2300°.

As will be seen later, it is important to calculate the minimum possible value of v, which would give rise to a visible meteor, since comparison with the observed minimum velocities of meteors will enable us to derive the temperature of the air.

In order that the meteor should be visible at a distance, D, the energy $4\pi D^2 E_{\text{min.}}$ must be radiated per second. If v is the velocity of the meteor, therefore, $dm/dt \,.\, v^2/2$ must be greater than $4\pi D^2 E_{\text{min.}}$ Hence $1/2\pi r^2 \,.\, dm/dt > 4D^2 E_{\text{min.}}/v^2 r^2$ the radius of the course, as shown above, being given by $(3/4\pi \,.\, m/\rho_m)^{1/3} = (6D^2 E_{\text{min.}} t_0/v^2\rho_m)^{1/3}$. This is the mass which must evaporate from unit surface in unit time in order that the meteor be visible at a distance D. If the material of which the meteor is composed is known, therefore, it defines the vapour pressure, and consequently the temperature, for, as shown above, it is equal to $\frac{1}{4}p'/V' = p'/4 \,.\, \sqrt{(3M'/RT_2)}$.

On the other hand, $T_2 < T_0 \,.\, (M_0 v^2/2RT_0)^{(\gamma-1)/\gamma}$. From these two expressions one can eliminate T_2, if p' is known as a function of T_2, and to obtain a minimum value for v. Strictly speaking, a correction should be added as the temperature of the cap exceeds the temperature of the meteor's surface by a calculable amount, but this is small and somewhat uncertain, since it involves a knowledge of the density of the air. It can only raise the final value, and as we are concerned with the minimum value it may be omitted.

Putting $E_{\text{min.}} = 10^{-7}$, $D = 10^7$, and taking iron, this leads to a value of p' of the order $10^4/t_0^{2/3}$, which means that T_2 must be at least 2000° if any reasonable value of t_0 is assumed. If T_0, the initial temperature of the air, is 220°, no meteors could therefore appear at a velocity less than $1{\cdot}9 \,.\, 10^6$ cm./sec. If T_0 is 300°, the minimum velocity would be about $1{\cdot}2 \,.\, 10^6$ cm./sec. Very similar values are obtained for olivine.

Height of Appearance of Meteors.

It is a matter of experience and of considerable interest that meteors in general appear comparatively suddenly, having passed through a fairly large mass of air without becoming luminous. The reasons for this are not far to seek. As we have seen, the first stages in which the molecules impinge individually on the meteor's surface are ineffective in causing appreciable evaporation or heating, since the shocks are so violent that the molecules and atoms are broken up and disturbed, and nearly all the energy is lost by radiation. The total energy lost is $\int \pi r^2 \frac{1}{2} (\rho v^3) dt$ from $\rho = 0$ to $\rho = 1{\cdot}4 . 10^{-10}/r$ for a meteor of velocity $4 . 10^6$ cm./sec., so that since r is small the meteor is invisible.

It is only when evaporation becomes considerable, and the cloud of vapour, with its comparatively large cross-section, is brought to rest by the air, that the luminosity becomes large and the meteor may be said to appear. Evaporation can only become appreciable when the surface temperature reaches some fairly high temperature T_2, and this cannot occur until a perceptible thickness has been heated to a similar if slightly lower temperature on account of heat conduction.

As has been shown, heat is supplied at the rate $k \frac{1}{2} (\rho v^3)$ per unit surface. It will be carried into the interior of the meteor at the rate $\kappa \, dT/dr$, if curvature is neglected. Hence heat will only become available for evaporation when dT/dr becomes of the same order as $\kappa/k \frac{1}{2} (\rho v^3)$. In the case of stone meteors κ is of the order $5 . 10^{-3}$ cal./degrees cm., so that since k is about $4{\cdot}2 . 10^{-3}$, we have $dT/dr = 10^{-8} \rho v^3$. For $\rho = 10^{-8}$ this is $3{\cdot}2 . 10^3$ degrees/cm. for a meteor moving at $4 . 10^6$ cm./sec. Hence, if its radius is $6 . 10^{-2}$ cm. as found above, the centre will be at almost the same temperature as the surface. For greater speeds and higher densities this would not be true, but fast meteors scarcely ever appear at low heights, *i.e.*, great densities. When they do they are usually very bright and evidently comparatively large objects. These have not been used in our calculations, for in these considerable complications are inevitable on account of the cold centre.

In general, then, in stone-, and *a fortiori* in iron-, meteors of small size, the surface can only begin to evaporate when the whole meteor has reached a temperature approaching that of the surface. Heat flows into the surface at a rate $\pi r^2 k \frac{1}{2} (\rho v^3)$, so that if m is the mass and s the specific heat, this point is defined by equating $\int ms \, dT_2$ to $\int \pi r^2 k \frac{1}{2} (\rho v^3) \, dt$. As $v dt = dl = dh \cos \chi$, therefore

$$msT_2 = \int \pi r^2 . k \tfrac{1}{2} (\rho v^3) \, dh/\cos \chi$$

Putting $\rho = \rho_0 \, e^{-g M_0/RT_0 . h}$ as in an isothermal atmosphere this reduces to

$$\frac{8 \, T_2 r \rho_m s \cos \chi}{3 v^2} = \int k \rho \, dh \text{ or } \rho_a = \frac{16}{3} \frac{T_2 r \rho_m s \cos \chi}{k v^2} . \frac{g M_0}{RT_0}, \qquad (2)$$

if the mean value of k is taken between the initial and final values of the surface temperature, and if the density above the height at which the cap forms is neglected. All the quantities on the right-hand side of the equation can be estimated with considerable accuracy since $r = (3m/4\pi\rho_m)^{1/3}$, and $m = 2Et_0/v^2$. T_0 is *a priori* doubtful, but may be derived by plotting ρ_a against the height. Therefore the density of the atmosphere at the height at which a meteor appears may be estimated with fair accuracy from the meteor's characteristics.

Height of Disappearance of Meteors.

Once the surface of the meteor has been raised to such a temperature that evaporation takes place in sufficient quantity to make the meteor luminous, practically all the heat which reaches the surface is used in volatilising the solid. The temperature is not high enough to cause appreciable dissociation or ionisation, and it is easily seen that radiation is small compared to the latent heat of the evaporated solid. In the case of iron, for instance, the total energy lost in the form of radiation at 2000° is less than a quarter of the heat used up in volatilisation. At higher temperatures this disparity is accentuated, since the radiation only increases with the fourth power of the temperature, whereas the expression for the vapour pressure contains the inverse temperature in the exponent.

As the heat reaching the surface is proportional to the density ρ_0, the rate of evaporation must be a measure of the density. Since deceleration can be neglected V_1 can be taken as constant, and as the vapour pressure varies rapidly with the temperature, changes in the surface temperature of the meteor, and consequently variations of V_2 can be left out of account. Hence the rate of evaporation is found as above by equating $l\,dm$ to $\pi r^2 . (V_1 - V_2)/3v . \frac{1}{2}(\rho v^3) . dt$, *i.e.*, the rate at which the radius decreases, is

$$dr/dt = \tfrac{1}{24}(V_1 - V_2)\, v^2/l \,.\, \rho_0/\rho_m.$$

If Δh is the observed difference in height between the meteor's appearance and disappearance, and L the length of path, one can put

$$v\,dt = L/\Delta h \,.\, dh,$$

or

$$dr = \tfrac{1}{24}(V_1 - V_2) \,.\, v/l \,.\, \rho_0/\rho_m \,.\, L/\Delta h \,.\, dh,$$

or

$$r_0 = \tfrac{1}{24}(V_1 - V_2) \,.\, vL/l\Delta h \,.\, 1/\rho_m \,.\, RT_0/g \,.\, \{\Sigma\, \rho_d/M_0 - \Sigma \rho_a/M_0\}, \qquad (3)$$

if $dp/dh = -\rho g M_0/RT_0$, as it would be in an isothermal atmosphere, and ρ_a and ρ_d denote the densities at which the meteor appears and disappears. In general, for any appreciable height drop ρ_a may be neglected compared to ρ_d, so that if a single gas only need be considered one finds

$$\rho_d = 24\, r_0/(V_1 - V_2) \,.\, l\,\Delta h/vL \,.\, \rho_m\, gM/RT_0, \qquad (4)$$

r_0, as has been shown, may be derived from the total energy, so that an estimate of ρ_d may be formed.

Using the value of k given in fig. 1, one finds, for instance, in the case of the typical meteor $\rho_d = 1{\cdot}05 . 10^{-7}$. These values are independent of those found from the heights of appearance. As will be seen they agree well with those calculated by the preceding formula.

Meteorites and Fireballs.

The above pages show that a consistent picture of meteoric phenomena is given by the theory outlined. It should be noted that it applies essentially to small meteors, *i.e.*, to those whose magnitude is less than that of Venus. Meteorites and fireballs, though they may be treated on similar lines, lead to different formulæ. They are of comparatively large mass, so that, *e.g.*, the assumption, justifiable in a particle of 1 mm. diameter, that the whole mass is nearly at the temperature of the surface, is no longer tenable. The temperature gradient at the surface is probably the same in both cases, but where the radius is measured in centimetres this involves a cold centre, whereas in a minute particle even a high gradient does not lead us to any such result. Again, meteorites penetrate to regions where the atmospheric density is comparatively high. If the velocity is $3 . 10^6$ cm./sec. and the density, *e.g.*, 10^{-5}, this will correspond to a blast of air at $4{\cdot}5 . 10^7$ dyne/cm., or 45 atmospheres. It is evident that at such pressures fine spray will be torn from the molten surface. Each drop will act as a miniature shooting star and a luminous trail will appear. If the meteorite is not round, the blast of air will cause it to rotate and the spray will spread in a cone. Seen from the side, this will appear as a double trail, whose angle is a measure of the ratio of peripheral speed to forward velocity. Incidentally, some such rotation, with the enormous centrifugal accelerations it involves, is the only means of explaining the fact that meteorites are sometimes observed to separate into two or more parts. Such separation could only be observed if the relative velocities were of the same order as the meteorite's velocity, and such high velocities can scarcely be explained in any other way.

For our purpose it is not necessary, however, to deal with fire-balls or meteorites, and we have confined ourselves to ordinary shooting stars in which complications such as the above do not occur, and whose large numbers enable inferences of value to be drawn from somewhat uncertain data.

Part II.—APPLICATION TO METEOROLOGY.

The main object of the reasoning set out above is to obtain equations by means of which the density of the atmosphere at great heights may be estimated from the characteristics of meteors occurring at these heights The only quantities which it is impossible to measure are the characteristics of the meteoric material itself, *e.g.*, the density, latent heat, and specific heat. Since the meteorites which fall to the ground consist, in varying proportions, of iron and stony substance, similar to olivine, we may fairly safely assume that the meteors have a similar composition. Except for this uncertainty, an accurate observation of a meteor—*e.g.*, by photography—would give a definite estimate of the density at its height. Such photographic observations are comparatively few, but it will be seen that much information of great meteorological interest may be obtained from the large number of eye observations of meteors which have been made.

Results obtained from Eye Observations of Meteors.

Mr. W. F. Denning has published particulars of several hundred observations of meteors, seen by himself and other observers. The values obtained from these meteors have been used to calculate the density of the atmosphere between 30 and 150 km., and also to find the temperature of the atmosphere at these heights, by using the equations (2) and (4).

It will be seen at once from the equations that it is important to know the velocity of the meteor accurately, since the density varies as the inverse cube of the velocity. Unfortunately, the eye observations of the velocity are likely to be very rough, and, as might be expected, the values obtained for the density at any height from different meteors are very discordant. We have, therefore, selected those meteors which were seen by the most skilled observers (as suggested by Mr. Denning), and the density at the height of appearance and disappearance has been calculated for each of these meteors.

As will be seen later, there is good evidence that the height of the meteor is obtained with very fair accuracy, even these eye observations having a probable error of not more than a few kilometres. In view of the accuracy obtained in determining the height, it is probable that the error in the length of path is not great. The brightness of the meteor must be subject to considerable inaccuracy, but fortunately it only appears as the one-third power. It has been assumed that the distance from the observer who estimates the brightness is equal to the average height, as this value is not given. The error involved will usually not be great. The error in the estimate of the velocity is obviously by far the most important.

It will be noticed that to obtain the density we require to know the temperature. We shall show that there are other reasons for thinking that the temperature above 60 km. is somewhere about 300° A., but it is obvious that if we calculate the densities with any assumed temperature, the rate of change of density with height will itself give the true temperature, and this is independent of the temperature assumed for the purpose of calculation.

We also require to know the molecular weight of the gases of the atmosphere, but up to at least 130 km. there is every reason to suppose that it is predominantly nitrogen and oxygen.

Density of Atmosphere Calculated from Heights of Appearance and Disappearance.

In fig. 3 the densities of the air as measured by "ballon sondes" up to 20 km. are shown by the continuous line AB. Above this height the broken

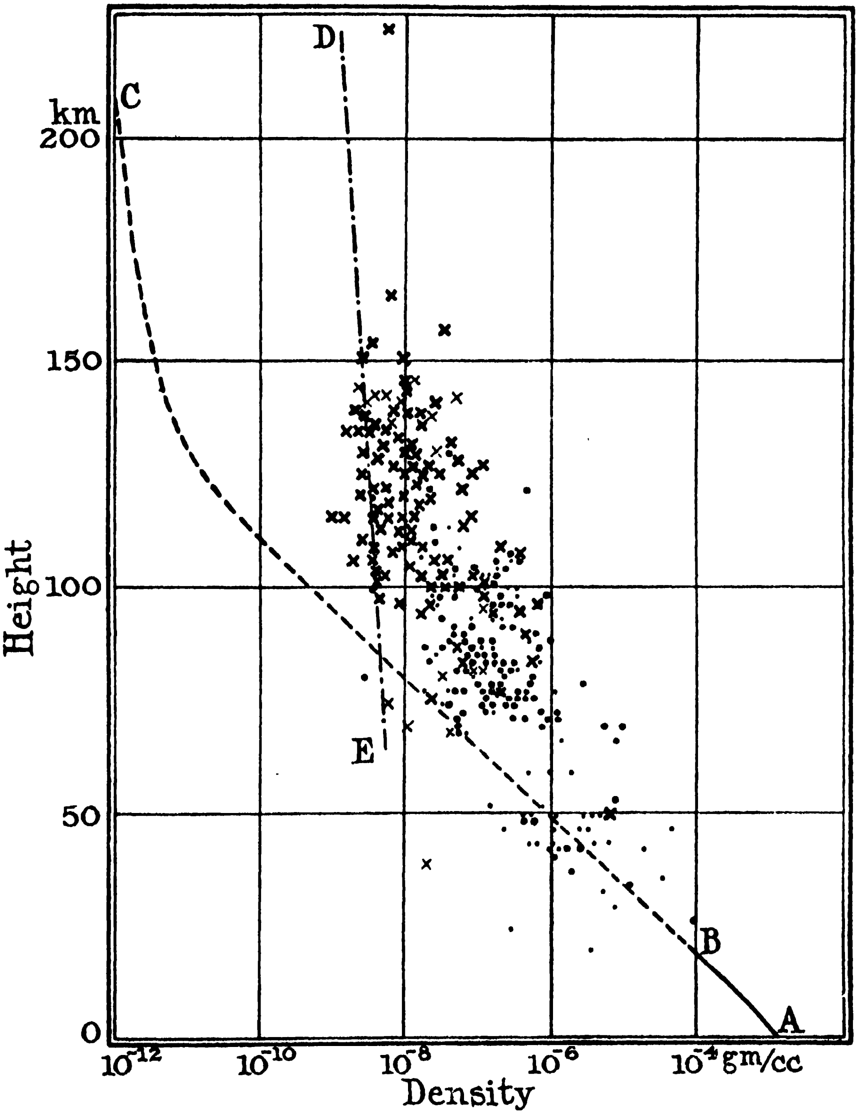


FIG. 3.

line BC indicates the densities calculated on the assumption that the whole upper atmosphere is isothermal at 220° A. Each dot represents the density of the air as calculated from the observations of one meteor for the height of disappearance of the meteor. Similarly the crosses represent the densities for the heights of appearance. The large dots and crosses refer to meteors seen by two of the selected observers, the small dots and crosses to meteors seen by one of these observers and one other (presumably less skilled) observer.

It will be seen that between 30 and 50 km. the densities obtained from meteor observations agree with those measured by "ballon sondes" up to 20 km. and the densities calculated on the assumption that the temperature continues at 220° A. at greater heights.

The number of meteors which disappear in this region is not great, but the evidence for a temperature of about 220° A. is quite good, and the usual assumption seems entirely justified.

Between 50 and 60 km. there is a gap in the observations as very few meteors either appear or disappear in this region. Between 60 and 160 km. there are abundant meteor observations, but here they all indicate densities very much greater than those calculated on the assumption of a uniform air temperature of 220° A. but consistent with a considerably higher temperature. Moreover, the rate of change of density with height shows that the temperature is much above 220° A.—probably about 300° A. Such a result is, of course, entirely contrary to previously accepted views on which it was supposed that the temperature would remain constant, or even decrease at very great heights. The difference between the density here calculated for, say, 150 km. and that previously assumed is of the order of 1000 times, and although the points on the diagram are not well grouped—probably owing to errors in estimated velocity—not one falls anywhere near the density which would occur if the temperature were 220° A. in the upper regions. One naturally wishes for confirmation from independent sources of a result so remarkable as this, and it is fortunate that three other arguments all lead to the same result, viz., a high temperature above 60 km.

It may be noted that the only likely source of error is that, in the larger meteors, the inner part may still be fairly cool when the outer part begins to vapourise. This would make the densities calculated for the height of appearance, too high. Such an error would undoubtedly occur in the case of very large fireballs, and the densities obtained from them are about ten times too great. These large meteors, however, have not been used in fig. 3, and the error in the small meteors cannot be large—possibly 50 per cent.—which is almost negligible compared to the differences of density obtained for temperatures of 220° A. and 300° A.

Minimum Density Calculated from Formation of Cap.

An independent check on the temperature above 60 km. can be obtained as follows:—An equation has been obtained above for the density at which a cap of compressed gas forms in front of the meteor. Before this cap forms, practically no heating occurs, hence the height at which this cap forms must be greater than that at which the meteor is first seen. Fortunately the quantities required to calculate the density at which the cap forms are all

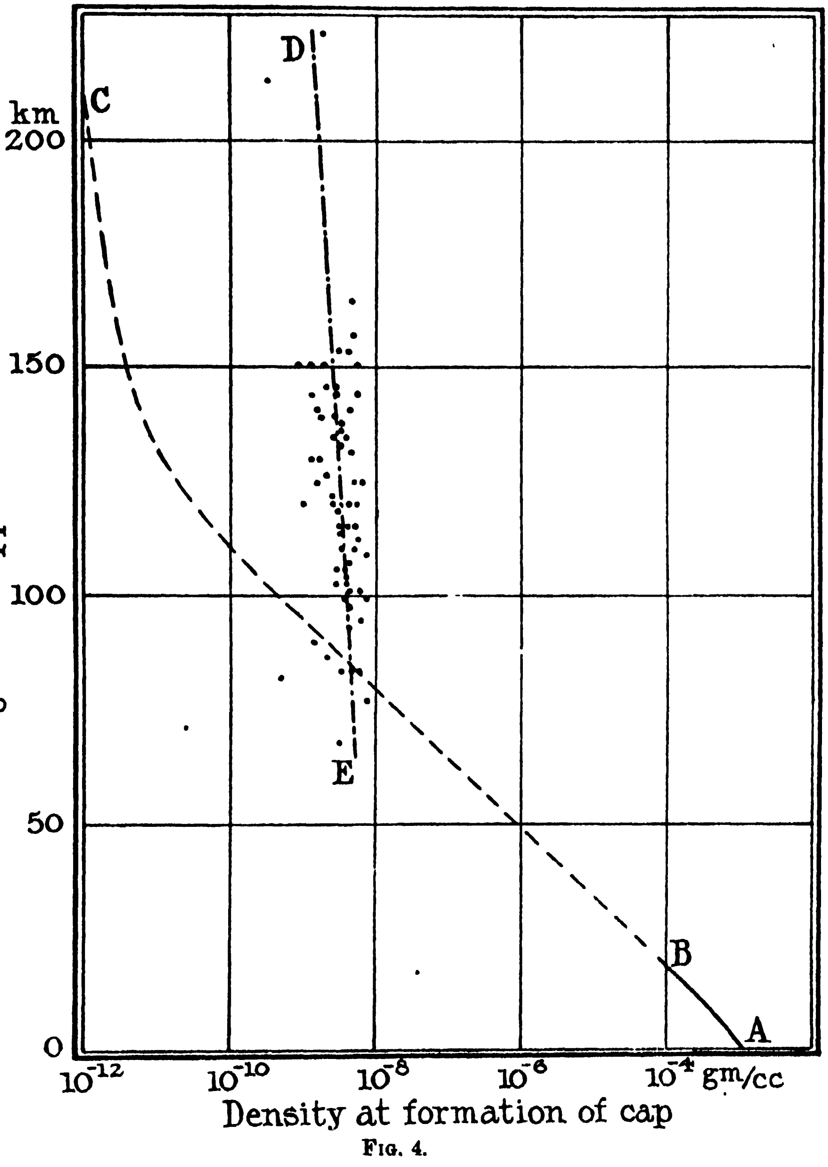


FIG. 4.

fairly accurately observed. Now while we do not know the height at which the cap formed, the density thus obtained will give a minimum possible value for the density at the height at which the meteor is first seen. In fig. 4 each dot represents the density calculated from one meteor at the height at which the cap was formed, plotted against the height of appearance of the meteor. The line ABC is the same as that in fig. 3. It is invariably found that the density is much above the value which would occur at the greater heights if the temperature of the air were 220° A. It is, however, quite consistent with the densities calculated as above from meteor observations. The dotted line DE is drawn through the centre of the group of dots and is reproduced again in fig. 3. It should be noted that the theory involved in calculating fig. 3 is quite independent of that involved in calculating fig. 4.

Explanation of Absence of Meteors Disappearing at about 55 *km.*

Another interesting point arises. It will be seen from fig. 6 that very few meteors disappear at heights between 50 km. and 60 km. On the supposition that the atmosphere is isothermal at 220° A in these regions, there seems to be no explanation for this curious effect. If, however, we suppose that the upper atmosphere has a much higher temperature, and that the change in temperature from about 220° A (which is known to be the temperature up to about 25 km.) occurs at about 60 km., then we might expect such an effect. As the meteor passes into colder air the temperature of the cap will fall, so that the heating will be reduced. As it proceeds further the increase in density will again cause sufficient heating to continue the rapid evaporation of the meteor. If this view is correct, and no other explanation of the phenomena has occurred to us, one might expect a diminution of brightness at the critical height. Observations in this region are therefore very desirable, for though quantitative examination shows an adequate effect for slow meteors it is small in the fast ones.

We may here point out that the relatively sharp limits of height to the region where few meteors disappear, is proof of the small probable error in the estimates of height by eye observations. Even supposing the limits were perfectly sharp, errors of observation would tend to smooth them out, and if the errors were large the minimum region would almost disappear. Moreover, in fig. 6, all the observations published by Mr. Denning are used, not only those by specially skilful observers. The inference is, then, that the probable error in the heights obtained by the best observers is very small, not more than a few kilometres.

Independent Check on Temperature from Minimum Velocity.

Another check is possible on the temperature of the atmosphere where oxidation of the meteor is a negligible source of heat. Since the meteor's surface must be heated to a certain high temperature before it gives off enough vapour to become visible, there will be a minimum velocity below which no meteor ought to occur, because the temperature of the cap in front of it will be too low. This minimum velocity will depend on the temperature of the air, and as shown above is 12 km./sec. for an air temperature of 300° A, and 19 km./sec. for an air temperature of 220° A. Unfortunately, eye observations of velocity are not very accurate, but there are a considerable number below 19 km./sec., again indicating a high temperature above 60 km. (see figs. 5 and 6). Lower down this method is invalid owing to possible oxidation.

Pressure and Density at the Auroral Level.

One of the reasons which make a knowledge of the density at great heights most important is the light it may throw on various theories put forward to account for auroræ. According to our results the density would be about 10^{-8} at 100 km., the lower limit of the auroræ, *i.e.*, about 100 times the density hitherto assumed. It might be possible to obtain an independent check on the density by comparing the observed widths of the vestigial nitrogen bands, which form most of the auroral lines, with the widths at known densities in the laboratory. Such a test might be feasible if the density is of the order of 10^{-8}, whereas, it would be hopeless if the density were 10^{-10}.

Examination of Assumptions.

Since the main result of this paper, *i.e.*, that the density, and consequently the temperature, of the atmosphere at great heights is very much higher than had previously been suspected, it behoves us to re-examine our initial assumptions with great care.

The first possibility that occurs to one is that the head-resistance may be less than $S\frac{1}{2}\rho v^2$, since some molecules might be supposed to escape before they had reached the meteor's velocity. In any case their number must be small, and the correction, if any, cannot be important. Such as it is, it would lead to values even further from those generally accepted, for if the head-resistance is less than $S\frac{1}{2}\rho v^2$, the true density must be greater than that calculated, assuming $F = S\frac{1}{2}\rho v^2$.

Again, the temperature derived by assuming adiabatic compression up to a pressure $\frac{1}{2}\rho v^2$ would be too high, so that here also any error in the initial formula only leads to a still more remarkable result.

The second possibility would be to deny the validity of the argument by which k was derived, and to say that a much larger fraction of the total work became available for heating the meteor. This would lead to a lower density, for if the whole, or a large fraction, of the energy were available for heating, one could assume a lower density, and still arrive at a sufficiently rapid rate of evaporation. This possibility is negatived, however, by the fact that our efficiency factor, k, does lead to correct results at low altitudes. If k were nearly unity it would be impossible for any meteors of ordinary

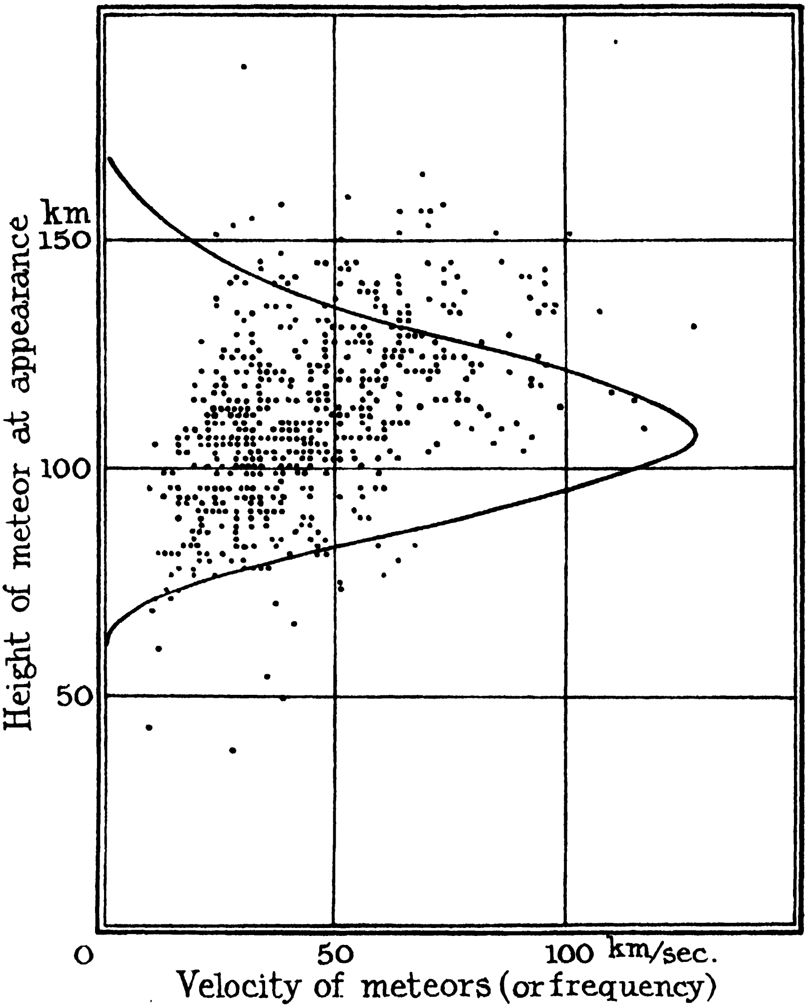


Fig. 5.

brightness to reach low altitudes, for the rate of evaporation, and consequently the brightness, would become enormous at relatively high densities. Again, whatever value one assigns to k, one cannot alter the slope of the density curve, and this is an independent measure of the temperature, which agrees with the value derived from the absolute density and with the figures arrived at from the observed minimum velocities.

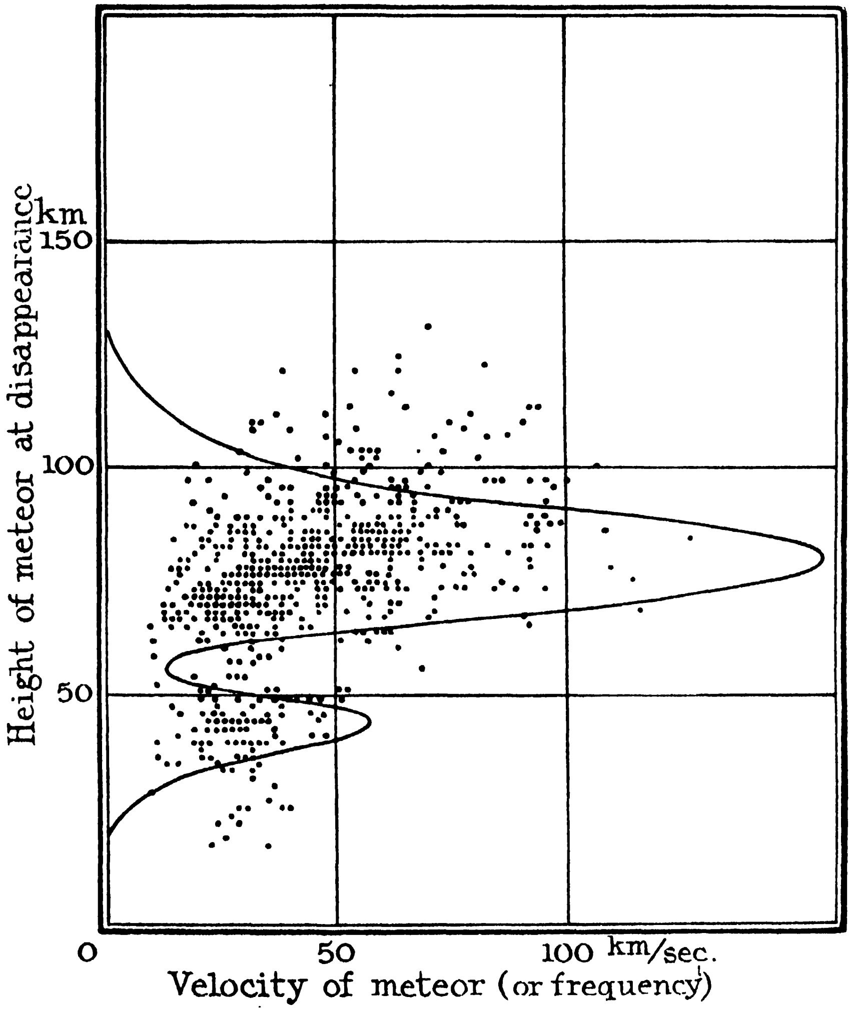


Fig. 6.

Finally, one might question the assumption that practically all the energy of the meteor is radiated as visible light. It may be that a fraction is not recorded in the ultra-violet and in the infra-red, and again some might be used up in other ways. The colour of meteors seems to indicate that we are not far wrong in assuming most of the radiation to occur in the visible part of the spectrum. If this were not so, meteors would appear violet or red instead of white. In any case, here again any correction would lead to a still higher temperature for the air. For if the energy were really underestimated, the mass of the meteor would also be higher than we have assumed, and this would involve more friction, that is a denser atmosphere, if the observed short duration is to be explained.

There does not seem to be any other fundamental hypothesis that can be questioned. Certain constants used may be uncertain, but in no case can an error of anything like the required amount be assumed. It seems, therefore, that the temperature at great heights is of the order of 280° to 300° abs.

Possible Explanation of Observed High Temperature.

It would lead too far here to try and do more than give an outline of the explanation of this main fact, namely, that the temperature at great heights rises to some 300°, whereas, as is well known, it amounts to but 220°, at from 10 to 20 km. For the sake of completeness, it may be worth while, however, to put forward briefly the sole hypothesis which has occurred to us that appears to be capable of explaining the observations, leaving a more detailed examination for another occasion.

The total absorption by the atmosphere above the region of convective equilibrium is small, probably not more than 20 per cent. of the total radiation. Hence, as a first approximation, one can calculate the radiative equilibrium temperature of a mass of gas, without reference to the layers above and below it, simply by equating the energy emitted by it to the energy absorbed. If κ_ν is the absorption coefficient of the gas at frequency ν, and ϵ_ν the emission coefficient of the earth, then calling the solid angle subtended by the earth 2π, and that of the sun, which we may take as a black body, α, we have as a condition of equilibrium

$$4\pi\int_0^\infty \frac{\kappa_\nu \nu^3}{e^{\beta\nu/T_0}-1}\,d\nu = 2\pi\int_0^\infty \frac{\epsilon_\nu\kappa_\nu\nu^3}{e^{\beta\nu/T_e}-1}\,d\nu + \alpha\int_0^\infty \frac{\kappa_\nu\nu^3}{e^{\beta\nu/T_s}-1}\,d\nu,$$

T_0 being the temperature of the gas, T_e that of the earth, and T_s that of the sun, and β being Planck's constant, h, divided by Boltzmann's constant, k.

Using this equation, it is possible to determine the radiative equilibrium temperature of any gas if ϵ_ν and κ_ν are known. As an example, we may take one of the simplest cases, CO_2, a gas in which κ_ν is negligible for all frequencies except those in the immediate neighbourhood of $2{\cdot}05 \, . \, 10^{13}$. Since the sun's radiation in this region is small, therefore $2/(e^{\beta\nu/T_0}-1) = \epsilon_\nu/(e^{\beta\nu/T_e}-1)$, whence putting $T_e = 280°$, one finds for $\epsilon_\nu = 1$, $T_0 = 236°$, and, *e.g.*, for $\epsilon_\nu = 1/2, T_0 = 206°$.

A more difficult case is that of H_2O, for which κ is extremely irregular. Again, the energy absorbed from the sun is small compared to that absorbed from the earth in the infra-red. Its absorption bands are widely distributed in the infra-red, and a rough estimate indicates that its equilibrium temperature would be in the neighbourhood of that of a grey body subject only to the earth's radiation, *i.e.*, $T_e/\sqrt[4]{2} = 237°$, for $\epsilon_\nu = 1$ and 200° for $\epsilon_\nu = 1/2$.

A somewhat different picture is presented by O_3. Here we have a single well-marked absorption band at frequency $3 \cdot 16 . 10^{13}$ and again strong absorption at frequencies above 10^{15}. In this case the sun's radiation, which is still considerable at frequencies above 10^{15}, cannot be neglected, and the equilibrium temperature amounts to about 280°.

The gases N_2, H_2, He have no well-marked absorption bands and those of O_2 are so narrow that the energy they take up is negligible. Hence these gases have an equilibrium temperature in the neighbourhood of that of a grey body, *i.e.*, 280°, but the tendency to take up this equilibrium temperature is small.

If now we consider the earth's atmosphere, which contains these gases in varying proportions, the mean equilibrium temperature must lie between these values. Lower in the stratosphere where there is a considerable amount of CO_2 and H_2O and some 20 per cent. of the earth's radiation is absorbed, the mean temperature will approach the temperature proper to these two gases, *i.e.*, it will be less than 236°, since the transparent gases are ineffective as radiators. Above 60 kilometres the CO_2 has practically disappeared and the H_2O is reduced to 1/100 of its amount at 10 km. Here, therefore, these gases only transmute some 0·2 per cent. of the total radiation and are almost as ineffective as N_2 or H_2. But at great heights O_3 becomes an important factor in the situation, for we know that O_3 is formed from O_2 by short-wave radiation (1600° A. to 1800° A.), and the greatest percentage of O_3 must therefore occur in the outermost layers of the atmosphere. O_3, as has been shown, has a comparatively high equilibrium temperature. In all it absorbs some 2 per cent. of the total radiation, so that it does not seem unlikely that its influence preponderates at heights above, say, 60 km., where the other gases absorb less than 0·2 per cent.

If this is so, it is at once evident that a high temperature, *i.e.*, one in the neighbourhood of 280°, is to be expected above 60 km., and this agrees with the results obtained above. This explanation is only put forward tentatively and awaits more careful examination. It is hoped, in another paper, to go into the question more fully, and perhaps to find graphically more accurate values for the mean temperature of gases mixed in varying proportions.

SUMMARY.

Formulæ are derived for the head-resistance and heating of a particle moving through the air at a speed great compared to that of sound.

It is shown that all major meteoric phenomena can be accounted for consistently, if the luminosity of the meteor is attributed to the collision of volatilised meteoric vapour with the air molecules.

From observed meteoric data the density or temperature of the air at great heights is derived in four independent ways, which give results consistent with each other.

The density above 60 km. appears to be very much greater than corresponds to an isothermal atmosphere at 220° abs., and the temperature appears to be in the neighbourhood of 300° abs. A tentative explanation is put forward to account for such a high temperature based on the radiative properties of ozone.

TABLE OF SYMBOLS

R	Gas constant.
g	Acceleration due to gravity.
N	Number of molecules per grm.-mol.
n_0	Number of molecules per c.c. at standard density.
L_0	Mean free path at standard density.
V_0	Velocity of molecules.
γ	Ratio of specific heats.
V	Velocity of sound.
ρ_0	Density of air at height of meteor.
ρ_a, ρ_d	Density of air at height of meteor's appearance and disappearance.
T_0	Temperature of air at height of meteor.
M_0	Molecular weight of air at height of meteor.
p_0	Atmospheric pressure at height of meteor.
r, r_0, r_1	Radius of meteor.
m	Mass of meteor.
ρ_m	Density of meteor.
v	Velocity of meteor.
S	Area of cross-section of meteor.
E	Energy radiated per second by meteor.
s	Specific heat of meteor.
l	Latent heat of meteor.
κ	Heat conductivity of meteor.
t_0	Duration of visible flight of meteor.
L	Length of visible path.
D	Distance of meteor from observer.
χ	Angle of meteor's path to vertical.
Δh	Difference of height of meteor at appearance and disappearance.
F	Head-resistance of meteor.
T_2	Temperature of meteor's surface.
M'	Molecular weight of vapour of meteor.

m'	Mass of air accelerated by meteor.
k	Efficiency factor for heating meteor.
p	Pressure of gas within cap in front of meteor $= F/S$.
T_1	Temperature of gas within cap.
p'	Vapour pressure of meteor at temperature T_2.
ν_1	Number of molecules per unit volume within cap having component towards meteor.
V_1	Velocity of molecules within cap having component towards meteor.
ν_2	Number of molecules per unit volume within cap having component away from meteor.
V_2	Velocity of molecules within cap having component away from meteor.
V'	Velocity of meteor relative to layer of gas immediately in front of it.
V_3	Velocity of gases leaving meteor.
T_v	A fictitious temperature $= M_0 v'^2/3R$.

The Theory of the Scattering of α- and β-Rays.

By J. H. JEANS, Sec.R.S.

(Received October 24, 1922.)

1. In what follows I have attempted to examine the theory of the passage of a beam of electrified particles through matter. The general theory is, naturally enough, highly complicated, but certain special cases are found to admit of fairly complete discussion. A summary of the results obtained will be found at the end (§ 13).

2. Consider a beam of N particles, which may be either positively or negatively charged. Suppose that, after passing through a thickness t of uniform scattering substance, the law of distribution of their directions of motion, as measured by co-ordinates θ, ϕ, is

$$N f(\theta, \phi, t)\, d\theta \frac{d\phi}{2\pi}. \tag{1}$$

Here θ is the angle which the direction of motion of the particle after traversing the thickness t of scattering substance makes with a certain standard line, while ϕ is an azimuth measured from any standard plane through this line. In the special case in which the beam was initially a parallel beam moving along the direction of the standard line $\theta = 0$, θ will, of course, measure the total deflection of the particle.

Since the whole number of particles in the beam is N, we obtain at once, on integration of formula (1),

$$\iint f(\theta, \phi, t)\, d\theta \frac{d\phi}{2\pi} = 1. \tag{2}$$

We shall be especially concerned with the law of distribution of deflections θ, apart from azimuths. In a beam of N particles, let this be

$$\mathrm{N}\mathrm{F}(\theta, t)\, d\theta, \tag{3}$$

where, of course,

$$\mathrm{F}(\theta, t) = \int f(\theta, \phi, t)\,\frac{d\phi}{2\pi} \text{ and } \int \mathrm{F}(\theta, t)\, d\theta = 1. \tag{4}$$

When f in any special problem happens to be independent of ϕ, it becomes identical with F.

3. Let us introduce temporarily (§§ 3–7) the special assumptions that the deflection produced by encounter with an atom is negligible when the nearest approach is greater than a certain quantity R, this quantity being assumed so small that the centres of adjacent atoms are never at a less distance than 2R. Then a moving particle cannot experience force from more than one atom at a time.

Consider a sheaf of particles, by definition

$$\mathrm{N} f(\theta_0, \phi_0, t)\, d\theta_0 \frac{d\phi_0}{2\pi} \tag{5}$$

in number, which, after having traversed a thickness t of foil, are moving in directions such that θ lies between θ_0 and $\theta_0 + d\theta_0$, while ϕ lies between ϕ_0 and $\phi_0 + d\phi_0$.

In traversing a further thickness dt of foil, these particles will experience encounters equal in number to

$$\mathrm{N}\pi\mathrm{R}^2 n\, dt \sec\theta_0 f(\theta_0, \phi_0, t)\, d\theta_0\, \frac{d\phi_0}{2\pi}, \tag{6}$$

where n is the number of atoms per unit volume, and the factor $\sec\theta_0$ enters because the path to be traversed by a particle which escapes encounters altogether is $dt \sec\theta_0$.

Let any new encounter produce a deflection θ_1, ϕ_1, where θ_1 is the angle between the directions of motion before and after the encounter, and ϕ_1 is

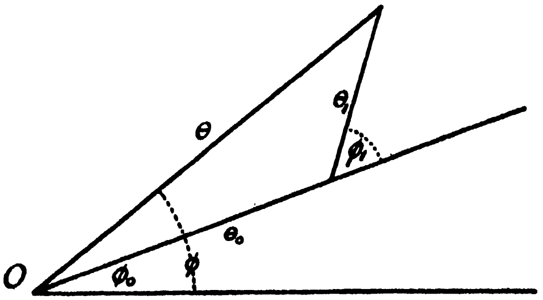


Fig. 1.

the change of azimuth measured through the direction of motion before the encounter. Suppose that the direction of motion after the encounter is specified by the angles θ, ϕ, these being measured in the same way as the angles θ_0, ϕ_0, which specified the direction of motion before the encounter. Then, by simple spherical trigonometry,

$$\cos\theta_1 = \cos\theta\cos\theta_0 + \sin\theta\sin\theta_0\cos(\phi-\phi_0). \tag{7}$$

As regards the deflection θ_1, ϕ_1 produced by this new encounter, it is clear that all values of ϕ_1 will be equally probable, while the law of distribution of values of θ_1 will not be dependent of the direction of motion before the encounter. Let us then assume for the law of distribution of θ_1, ϕ_1

$$\chi(\cos\theta_1)\sin\theta_1\, d\theta_1 \frac{d\phi_1}{2\pi}, \tag{8}$$

where χ depends only on the properties of the atom and of the particle, and is, of course, such that

$$\int \chi(\cos\theta_1)\sin\theta_1\, d\theta_1 = 1. \tag{9}$$

Out of the original beam of N particles, the number for which θ_0, ϕ_0, θ_1, ϕ_1 lie within a range $d\theta_0 d\phi_0 d\theta_1 d\phi_1$ will be

$$\mathrm{N}\pi\mathrm{R}^2 n\,dt \sec\theta_0 f(\theta_0, \phi_0, t)\,\chi(\cos\theta_1)\sin\theta_1 d\theta_1 \frac{d\phi_1}{2\pi}\, d\theta_0 \frac{d\phi_0}{2\pi}. \tag{10}$$

The final deflection of these particles, after passing through the thickness $t+dt$ of foil, will be θ, determined by equation (7). Let us accordingly change our variables in expression (10) from θ_1, ϕ_1 to θ, ϕ. The differentials are connected by the relation

$$\sin\theta_1 d\theta_1 d\phi_1 = \sin\theta d\theta d\phi,$$

since each side of this equation represents an element of area on a unit sphere. Consequently, expression (10) transforms into

$$\mathrm{N}\pi\mathrm{R}^2 n\,dt \sec\theta_0 f(\theta_0, \phi_0, t)\,\chi(\cos\theta_1)\sin\theta d\theta \frac{d\phi}{2\pi}\, d\theta_0 \frac{d\phi_0}{2\pi}, \tag{11}$$

in which $\cos\theta_1$ is given in terms of the new variables by equation (7). If we integrate this expression over all possible values of θ_0, ϕ_0 and ϕ, we shall obtain the total number of particles which move in directions for which θ lies between θ and $\theta+d\theta$, after experiencing an encounter in the thickness dt of foil, and, of course, after experiencing any number of encounters (including zero) in the previous thickness, t, of foil.

One integration can be carried out at once. Write ψ for $\phi-\phi_0$, so that, in place of (7), we have

$$\cos\theta_1 = \cos\theta\cos\theta_0 + \sin\theta\sin\theta_0\cos\psi, \tag{12}$$

and change variables from θ, ϕ, θ_0, ϕ_0 to θ, ψ, θ_0, ϕ_0 in expression (11). If

$\cos\theta_1$ is understood to be given by equation (12), the whole change is effected by replacing $d\phi$ by $d\psi$. It is now possible to integrate expression (11) with respect to ϕ_0. The result (*cf.* equation (4)) is

$$N\pi R^2 n\,dt \sec\theta_0 F(\theta_0, t)\,\chi(\cos\theta_1)\sin\theta\,d\theta\,\frac{d\theta_0\,d\psi}{2\pi}. \tag{13}$$

This expression, integrated over all values of θ_0 and ψ, gives the number of which we are in search.

4. By definition of F, the number of particles which move in directions making angles between θ and $\theta+d\theta$ with the initial line, after passing through the thickness $t+dt$ of foil, is

$$NF(\theta, t+dt)\,d\theta.$$

This number may be regarded as the sum of two contributions, one arising from those particles which experienced an encounter in the last layer of foil of thickness dt, and the other arising from those particles which traversed this layer without an encounter, and so must have had values of θ_0 intermediate between θ and $\theta+d\theta$ after traversing the thickness t of foil.

The first contribution is precisely the number we calculated in the last section, while the second contribution is obviously

$$N(1-\pi R^2 n\,dt\sec\theta)\,F(\theta, t)\,d\theta.$$

Combining the two contributions, we must have

$$NF(\theta, t+dt)\,d\theta = N\pi R^2 n\,dt\sin\theta\,d\theta\iint F(\theta_0, t)\,\chi(\cos\theta_1)\frac{d\theta_0\,d\psi}{2\pi\cos\theta_0}$$
$$+N(1-\pi R^2 n\,dt\sec\theta)\,F(\theta, t)\,d\theta.$$

The left-hand member of this equation is equal to

$$NF(\theta, t)\,d\theta+N\frac{\partial F}{\partial t}\,dt\,d\theta,$$

so that the equation can be rewritten in the form

$$\frac{\partial}{\partial t}F(\theta, t)+\pi R^2 n\sec\theta\,F(\theta, t)$$
$$=\pi R^2 n\sin\theta\iint F(\theta_0, t)\,\chi[\cos\theta\cos\theta_0+\sin\theta\sin\theta_0\cos\psi]\frac{d\theta_0\,d\psi}{2\pi\cos\theta_0}. \tag{14}$$

5. In the majority of experimental researches on scattering, the maximum deflection studied has been only a fraction of a radian. This suggests the suitability of an expansion in powers of θ_0. We put

$$\cos\theta_0 = 1-\tfrac{1}{2}\theta_0^2+\tfrac{1}{24}\theta_0^4-\ldots,$$
$$\sec\theta_0 = 1+\tfrac{1}{2}\theta_0^2+\tfrac{5}{24}\theta_0^4+\ldots,$$

and obtain

$$\frac{\chi[\cos\theta\cos\theta_0+\sin\theta\sin\theta_0\cos\psi]}{\cos\theta_0}$$

$$\begin{aligned}&=\chi(\cos\theta)+\tfrac{1}{2}\theta_0^2\{\chi(\cos\theta)-\cos\theta\,\chi'(\cos\theta)+\sin^2\theta\cos^2\psi\,\chi''(\cos\theta)\}\\&\quad+\theta_0^4\{\tfrac{5}{24}\chi(\cos\theta)-\tfrac{5}{24}\cos\theta\,\chi'(\cos\theta)+(\tfrac{1}{8}\cos^2\theta+\tfrac{11}{48}\sin^2\theta\cos^2\psi)\chi''(\cos\theta)\\&\qquad-\tfrac{1}{4}\cos\theta\sin^2\theta\cos^2\psi\,\chi'''(\cos\theta)+\tfrac{1}{24}\sin^4\theta\cos^4\psi\,\chi''''(\cos\theta)\}\\&\quad+\text{terms in }\theta_0,\ \theta_0^3,\ \theta_0^5,\ \theta_0^6,\text{ etc.}\end{aligned}$$

We substitute this value in the integral on the right-hand of equation (14), and carry out the integration with respect to ψ. The terms in θ_0, θ_0^3, θ_0^5 ... all vanish on integration, and, on making use of the relation

$$\int F(\theta_0, t)\,d\theta_0=1,$$

the equation assumes the form

$$\begin{aligned}&\frac{\partial}{\partial t}F(\theta, t)+\pi R^2n\sec\theta\,F(\theta, t)\\&\qquad=\pi R^2n\sin\theta\,[\chi(\cos\theta)\\&\qquad\quad+\tfrac{1}{2}\{\chi(\cos\theta)-\cos\theta\,\chi'(\cos\theta)+\tfrac{1}{2}\sin^2\theta\,\chi''(\cos\theta)\}\int F(\theta_0, t)\,\theta_0^2 d\theta\\&\qquad\quad+\{\tfrac{5}{24}\chi(\cos\theta)-\tfrac{5}{24}\cos\theta\,\chi'(\cos\theta)+(\tfrac{1}{8}\cos^2\theta+\tfrac{11}{96}\sin^2\theta)\chi''(\cos\theta)\\&\qquad\qquad-\tfrac{1}{8}\cos\theta\sin^2\theta\,\chi'''(\cos\theta)+\tfrac{1}{64}\sin^4\theta\,\chi''''(\cos\theta)\}\int F(\theta_0, t)\,\theta_0^4 d\theta\\&\qquad\quad+\text{ terms in }\theta_0^6,\text{ etc.}]\end{aligned}\tag{15}$$

6. To solve, let us assume for $F(\theta, t)$ an expansion in powers of t in the form

$$F(\theta, t)=\Phi_0(\theta)+t\Phi_1(\theta)+\tfrac{1}{2}t^2\Phi_2(\theta)+\ldots\tag{16}$$

so that we may put, in equation (15),

$$\int F(\theta_0, t)\,\theta_0^2 d\theta_0=\int\Phi_0(\theta_0)\,\theta_0^2 d\theta_0+t\int\Phi_1(\theta_0)\,\theta_0^2 d\theta_0+\tfrac{1}{2}t^2\int\Phi_2(\theta_0)\,\theta_0^2 d\theta_0+\ldots\tag{17}$$

and similarly for $\int F(\theta_0, t)\,\theta_0^4 d\theta_0$, etc.

It will be convenient at this stage to introduce the further assumption that the incident beam is nearly parallel to the axis $\theta=0$. To be more precise we shall assume that $F(\theta, 0)$, which is the same thing as $\Phi_0(\theta)$, vanishes except when θ is so small that θ^2 may be neglected. We may now take

$$\int\Phi_0(\theta_0)\theta_0^2 d\theta_0=\int\Phi_0(\theta_0)\theta_0^4 d\theta_0=\ldots=0,$$

and equation (17) becomes

$$\int F(\theta_0, t)\theta_0^2 d\theta_0=t\int\Phi_1(\theta_0)\theta_0^2 d\theta_0+\tfrac{1}{2}t^2\int\Phi_2(\theta_0)\theta_0^2 d\theta_0+\ldots,$$

and similarly for $\int F(\theta_0, t)\theta_0^4\,d\theta_0$, etc.

To examine the error involved in this assumption, we may notice that if the incident beam is spread over the surface of a cone of angle 12° (semi-angle 6°), the value of $\int \Phi_0(\theta_0)\,\theta_0^2 d\theta_0$ is only 0·0109, representing an error of about 1 per cent. If the beam were spread throughout the same cone the error would, of course, be less. At present the error of all experimental measurements is probably well over 1 per cent. So that our assumption may be regarded as legitimate provided the angle of the cone of incidence is less than about 12°.

On substituting the assumed solution (16) into equation (15), and equating the coefficients of various powers of t, we obtain

$$\Phi_1(\theta) + \pi R^2 n \Phi_0(\theta) \sec\theta = \pi R^2 n \chi(\cos\theta) \sin\theta, \tag{18}$$

$$\begin{aligned} &\Phi_2(\theta) + \pi R^2 n \Phi_1(\theta) \sec\theta \\ &\quad = \tfrac{1}{2}\pi R^2 n \sin\theta \{\chi(\cos\theta) - \cos\theta\, \chi'(\cos\theta) + \tfrac{1}{2}\sin^2\theta\, \chi''(\cos\theta)\} \int \Phi_1(\theta_0)\theta_0^2 d\theta_0 \\ &\qquad + \pi R^2 n \sin\theta \{\tfrac{5}{24}\chi(\cos\theta) - \ldots \qquad\} \int \Phi_1(\theta_0)\theta_0^4 d\theta_0 + \ldots, \end{aligned} \tag{19}$$

and so on. These equations give the values of Φ_1, $\Phi_2 \ldots$ in succession, and with their help, equation (16) gives the law of distribution of deflections in the form

$$\begin{aligned} F(\theta, t) = F(\theta, 0) &+ \pi R^2 nt[\chi(\cos\theta)\sin\theta - F(\theta, 0)\sec\theta] \\ &+ \tfrac{1}{2}(\pi R^2 nt)^2 [G + F(\theta, 0)\sec^2\theta] + \ldots \end{aligned} \tag{20}$$

where

$$\begin{aligned} G = &-\chi(\cos\theta)\tan\theta \\ &+ \tfrac{1}{2}\sin\theta\,\{\chi(\cos\theta) - \cos\theta\,\chi'(\cos\theta) + \tfrac{1}{2}\sin^2\theta\,\chi''(\cos\theta)\} \int \chi(\cos\theta_0)\sin\theta_0\,\theta_0^2 d\theta_0 \\ &+ \sin\theta\{\tfrac{5}{24}\chi(\cos\theta) - \ldots \qquad\} \int \chi(\cos\theta_0)\sin\theta_0\,\theta_0^4 d\theta_0 + \ldots \end{aligned} \tag{21}$$

7. The primary object of experiments on the scattering of α- and β-rays is to determine or verify the atomic law of force, and this can be found at once if the function $\chi(\cos\theta)$ can be determined. Thus it is at once clear that the experimental material of greatest value is the determination of the coefficient of t in equation (20), or what is the same thing, of the coefficient ϵ in the law

$$\frac{I}{I_0} = 1 - \epsilon t, \tag{22}$$

which can be deduced from equation (20) when the terms in t^2 are neglected.

The intensity I of the beam which passes, after scattering, through a stop of semi-angle θ greater than that of the incident beam is

$$I = \int_0^\theta F(\theta, t)\, d\theta,$$

and if we use the value of $F(\theta, t)$ given by equation (20), with terms in t^2 neglected, this becomes

$$\frac{\mathrm{I}}{\mathrm{I}_0} = 1 - \pi\mathrm{R}^2 nt\left[1 - \int_0^\theta \chi(\cos\theta)\sin\theta\,d\theta\right]. \tag{23}$$

It will be remembered that $\chi(\cos\theta)\sin\theta d\theta$ has been taken to represent the distribution of deflections produced by the passage of a particle through a sphere of radius R surrounding an atomic nucleus. Let an orbit in which the perpendicular from the nucleus on to the orbit at its first entry into the sphere is p show a total deflection θ. The law of distribution of values of p is obviously

$$\frac{2p\,dp}{\mathrm{R}^2},$$

while that of values of θ is $\chi(\cos\theta)\sin\theta\,d\theta$. Thus we must have

$$\frac{2p\,dp}{\mathrm{R}^2} = \chi(\cos\theta)\sin\theta\,d\theta,$$

giving on integration

$$\int_0^\theta \chi(\cos\theta)\sin\theta\,d\theta = \int_p^{\mathrm{R}} \frac{2p\,dp}{\mathrm{R}^2} = \frac{\mathrm{R}^2 - p^2}{\mathrm{R}^2}.$$

Equation (23) now reduces to

$$\frac{\mathrm{I}}{\mathrm{I}_0} = 1 - \pi p^2 nt, \tag{24}$$

so that if ϵ can be determined experimentally, we can at once deduce the value of p corresponding to any value of θ.

For instance, if the relativity correction is neglected altogether, and the nucleus attracts according to the law of the inverse-square, it is readily found from ordinary particle dynamics, that

$$p^2 = \frac{\mathrm{R}^2}{1 + c^2\tan^2(\frac{1}{2}\theta)},$$

where $c = 1 + \frac{2\mathrm{R}}{b}$ and $b = \frac{2\mathrm{E}e}{mv^2}$ in the ordinary notation.

If R is large compared with b, this reduces to

$$p^2 = 4b^2\cot^2(\tfrac{1}{2}\theta),$$

and equation (24) gives at once the familiar law of single-scattering.

8. We have now found that, subject to the assumptions we have made, the law of scattering will be of the form

$$\frac{\mathrm{I}}{\mathrm{I}_0} = 1 - \epsilon t$$

for small values of t, and the coefficient ϵ will be exactly that calculated on the hypothesis of single-scattering. Our assumptions have been that the fields of force of neighbouring atoms do not overlap, and that the angle of the incident beam is small (in practice a semi-angle of not more than about 6°).

Our result that a small divergence of the incident beam produces no appreciable influence on the total intensity of the beam passing through a stop of angle greater than that of the incident beam is at variance with the statement of Schonland,* that a divergence α of the incident beam will increase the scattering in a ratio γ given by

$$\gamma = \frac{\sec^2 \phi}{(1-\alpha^2/\phi^2)^2},$$

where ϕ is the value of θ for the scattered beam. No information is given as to the way in which this formula is reached, but a few test cases immediately show it to be erroneous. For instance,

when $\alpha = 0$, true value is $\gamma = 1$; formula gives $\gamma = \sec^2 \phi$

when $\phi = \alpha$ and $t = 0$, true value is $\gamma = 1$; formula gives $\gamma = \infty$.

Thus, in spite of Schonland's statement to the contrary, I still consider it proved that if the fields of force of separate atoms may be assumed not to overlap, then the observed scattering ought to be that given by the simple law of single-scattering up to angles of incidence which are well above those which have so far been used in experiments on β-rays, provided only that we do not go beyond the range of thickness of scattering material within which I/I_0 is a linear function of t. But before drawing any inferences from this result, we must examine the effect of overlap, since some must undoubtedly occur, in the atomic fields of force.

Let us divide the field of force of each atom into two parts—an internal part, of extent so small that these parts never overlap, and an external part, in which the fields are assumed to be so feeble that the deflection produced by the external field of a single atom on a moving β-particle may be treated as a small quantity. Let the surface of demarcation between these two fields be a sphere of radius R surrounding the nucleus of each atom. Then the field of force of all the atoms can be divided into two parts:—

(*a*) A field which exists only inside the spheres of radius R, such that the field inside any one sphere arises solely from the nucleus at the centre of that sphere.

(*b*) The remainder of the field which is nowhere very intense and is made up of contributions, each of feeble amount, from the separate atoms.

The deflection of a β-particle produced by field (*a*) will obey the laws we have just investigated, but to obtain the true deflection, we must compound this deflection with that produced by field (*b*). Now this latter deflection is precisely of the type considered by Sir J. J. Thomson when he formulated his original theory of "multiple-scattering," and will therefore obey the law of multiple-scattering.

---

* 'Roy. Soc. Proc.,' A, vol. 101, p. 299 (1922).

9. Out of a beam of N particles, the number which experience a deflection between θ_0 and $\theta_0+d\theta_0$ due to field (*a*), and a deflection between θ_1 and $\theta_1+d\theta_1$ due to field (*b*), this latter deflection making an angle between ϕ and $\phi+d\phi$ with the former, will be

$$\mathrm{NF}(\theta_0, t)\, d\theta_0\, e^{-\theta_1^2/\kappa t}\, d\left(\frac{\theta_1^2}{\kappa t}\right)\left(\frac{d\phi}{2\pi}\right), \tag{25}$$

where κ is a constant of the scattering material such that $\mathrm{I}/\mathrm{I}_0 = 1-e^{-\theta_1^2/\kappa t}$ represents the law of multiple-scattering.

The total deflection θ of such a particle is, of course, given by

$$\cos\theta = \cos\theta_0 \cos\theta_1 - \sin\theta_0 \sin\theta_1 \cos\phi. \tag{26}$$

Unfortunately it does not prove possible to obtain the law of distribution of values of θ in a form which is simple enough to be of any practical use, except in certain special cases.

It will be remembered that Thomson's investigation of multiple-scattering proceeds on the supposition that θ_1^4 may be neglected. Let us now assume also that θ_0^4 may be neglected, *i.e.*, we definitely limit ourselves to such small angles of scattering that their fourth powers may be neglected (if $\theta = 30^0$, $\theta^4 = 0{\cdot}075$). Then relation (26) may be replaced by

$$\theta^2 = \theta_0^2 + \theta_1^2 + 2\theta_0\,\theta_1 \cos\phi, \tag{27}$$

and the deflections θ_0, θ_1 now follow the same law of composition as vectors in a plane. Let ψ, η denote the two components of θ_1 along and perpendicular to the direction of θ_0. Then

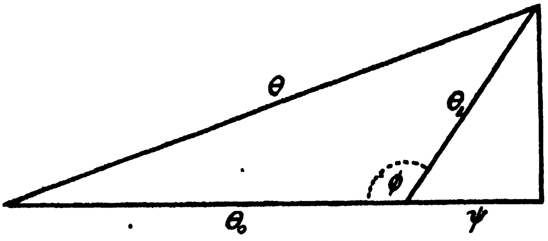


FIG. 2.

$$\theta_1^2 = \psi^2 + \eta^2 \qquad \text{and} \qquad \theta_1\, d\theta_1\, d\phi = d\psi\, d\eta.$$

The law of distribution (25) may now be expressed in the form

$$\mathrm{NF}(\theta_0, t)\, d\theta_0\, \frac{1}{\pi\kappa t}\, e^{-\psi^2/\kappa t}\, d\psi\, e^{-\eta^2/\kappa t}\, d\eta. \tag{28}$$

Let ω denote the total deflection $\theta_0+\psi$ in the direction of θ_0, then

$$\begin{aligned}\mathrm{F}(\theta_0, t) &= \mathrm{F}(\omega-\psi, t)\\ &= \mathrm{F}(\omega, t) - \psi\frac{\partial \mathrm{F}}{\partial\omega} + \tfrac{1}{2}\psi^2\frac{\partial^2 \mathrm{F}}{\partial\omega^2} - \tfrac{1}{6}\psi^3\frac{\partial^3 \mathrm{F}}{\partial\omega^3} + \tfrac{1}{24}\psi^4\frac{\partial^4 \mathrm{F}}{\partial\omega^4} - \dots\end{aligned} \tag{29}$$

On changing variables from θ_0, ψ, η to ω, ψ, η in expression (28), we may replace $d\theta_0$ by $d\omega$, and on using the value for $F(\theta_0, t)$ just found in equation (29) and integrating with respect to all values of ψ, the law assumes the form

$$N\left[F(\omega, t)+\tfrac{1}{4}\kappa t\frac{\partial^2 F(\omega, t)}{\partial\omega^2}+\tfrac{1}{64}(\kappa t)^2\frac{\partial^4 F(\omega, t)}{\partial\omega^4}+\dots\right]e^{-\eta^2/\kappa t}\frac{\partial\omega\,\partial\eta}{\sqrt{(\pi\kappa t)}}. \quad (30)$$

The total deflection θ is now given by $\theta^2 = \omega^2+\eta^2$, and we proceed to change the variables from ω, η to θ, η. The differentials are connected by

$$\omega\, d\omega\, d\eta = \theta\, d\theta\, d\eta. \quad (31)$$

By definition $F(\omega, t)\,d\omega$ is the proportion of particles which are scattered by field (a) so as to be within a ring bounded by cones of angles ω and $\omega+d\omega$. It follows that, near the value $\omega = 0$, $F(\omega, t)$ must be of the form $\omega\times$ a function of ω^2—*i.e.*, it must be an odd function of ω. Thus we may write

$$F(\omega, t)+\tfrac{1}{4}\kappa t\frac{\partial^2 F(\omega, t)}{\partial\omega^2}+\tfrac{1}{64}(\kappa t)^2\frac{\partial^4 F(\omega, t)}{\partial\omega^4}+\dots = \omega\Phi(\omega^2), \quad (32)$$

and the law of distribution (25) becomes

$$N\Phi(\omega^2)\,e^{-\eta^2/\kappa t}\frac{\omega\, d\omega\, d\eta}{\sqrt{(\pi\kappa t)}}. \quad (33)$$

Since $\omega^2 = \theta^2-\eta^2$, we may expand $\Phi(\omega^2)$ in the form

$$\Phi(\omega^2) = \Phi(\theta^2)-\eta^2\frac{\partial\Phi(\theta^2)}{\partial\theta^2}+\tfrac{1}{2}\eta^4\frac{\partial^2\Phi(\theta^2)}{(\partial\theta^2)^2}-\dots \quad (34)$$

On substituting this value for $\Phi(\omega^2)$ into equation (33) and replacing $\omega\, d\omega\, d\eta$ by $\theta\, d\theta\, d\eta$ (equation 31), we obtain the law of distribution in terms of θ and η. It is immediately possible to integrate with respect to η, and we find the law of distribution with respect to θ to be

$$N\left[\Phi(\theta^2)-\tfrac{1}{2}\kappa t\frac{\partial\Phi(\theta^2)}{\partial\theta^2}+\tfrac{3}{8}(\kappa t)^2\frac{\partial^2\Phi(\theta^2)}{(\partial\theta^2)^2}-\dots\right]\theta\, d\theta. \quad (35)$$

On inserting the value of $\Phi(\theta^2)$ from equation (32), we obtain the law of distribution expanded in powers of κt.

10. If we neglect κt altogether, the law (35) reduces to

$$N\theta\Phi(\theta^2)d\theta = NF(\theta, t)\, d\theta,$$

which is, of course, the law of scattering when field (b) is neglected entirely.

If we retain first powers only of κt, neglecting squares, the law of distribution becomes

$$\begin{aligned} &N\left[\theta\Phi(\theta^2)-\tfrac{1}{4}\kappa t\frac{\partial}{\partial\theta}\Phi(\theta^2)\right]d\theta \\ &\quad= N\left[F(\theta, t)+\tfrac{1}{4}\kappa t\frac{\partial^2 F(\theta, t)}{\partial\theta^2}-\tfrac{1}{4}\kappa t\frac{\partial}{\partial\theta}\left(\frac{1}{\theta}F(\theta, t)\right)\right]d\theta \\ &\quad= N\left[F(\theta, t)+\tfrac{1}{4}\kappa t\frac{\partial}{\partial\theta}\left\{\theta\frac{\partial}{\partial\theta}\left(\frac{1}{\theta}F(\theta, t)\right)\right\}\right]d\theta. \quad (36)\end{aligned}$$

Integrating with respect to θ from 0 to θ, we obtain

$$\frac{I}{I_0}=\int_0^\theta\left[F(\theta,t)+\tfrac{1}{4}\kappa t\frac{\partial}{\partial\theta}\left\{\theta\frac{\partial}{\partial\theta}\left(\frac{1}{\theta}F(\theta,t)\right)\right\}\right]d\theta$$

$$=\int_0^\theta F(\theta,t)\,d\theta+\tfrac{1}{4}\kappa t\theta\frac{\partial}{\partial\theta}\left(\frac{1}{\theta}F(\theta,t)\right), \qquad (37)$$

there being no additional term arising from the lower limit of integration, since clearly $1/\theta\,F(\theta,t)$ will be either a maximum or a minimum at $\theta=0$. In general, this quantity will decrease as we pass outwards from $\theta=0$, so that the last term in equation (37), which represents the effect of multiple scattering, is negative.

11. Equation (37) gives the expansion of the distribution function as far as the first power of κt. We may, however, inquire what is the expansion as far as the first power of t; that is to say, over the range for which I/I_0 is a linear function of t, what is the form of this function ?

To obtain the expansion as far as first powers of t, we expand $F(\theta,t)$ as far as first powers of t in the integral on the right-hand of equation (37), and replace $F(\theta,t)$ by $F(\theta,0)$ in the remaining term. Now if α is the semi-angle of the incident beam, $F(\theta,0)$ is equal to zero as soon as θ exceeds α. Thus, provided only that the angle of the stop is greater than that of the incident beam, equation (37) reduces to

$$\frac{I}{I_0}=\int_0^\theta F(\theta,t)\,d\theta, \qquad (38)$$

showing that the effect of multiple-scattering has disappeared altogether.

Thus throughout the range, if such exist at all, in which I/I_0 obeys a law of the form

$$\frac{I}{I_0}=1-\epsilon t,$$

the value of ϵ is precisely that which would be calculated on the hypothesis of single scattering.

12. The expansion as far as first powers of t, which we have just discussed, is invested with a certain interest by the circumstance that Crowther and Schonland believed that they had found a very substantial range, in which the experimental values indicated that I/I_0 varied in a linear way with t. Had it not been for this circumstance, an expansion as far as first powers of t would have been of but little interest, for the following reason :—

In expanding the functions F and Φ by Taylor's theorem, we have assumed their differential coefficients to be finite and continuous throughout the range under consideration. In actual practice, $F(\theta,0)$ is likely to change very abruptly near the boundary of the incident beam. We may undoubtedly represent this function as closely as we please by a function which, with its differential coefficients, will be finite and continuous, but, the nearer we approximate to the truth, the larger will these differential coefficients become. There may be but little practical value in an expansion in which t^2 is treated as a small quantity, and neglected when it is liable to be multiplied by a very large differential coefficient.

For this reason it is better, in actual practice, to treat the law of distribution (28) somewhat differently. A glance at equation (20) shows that the value of $F(\theta, t)$ is the sum of two parts, one depending only on $F(\theta, 0)$ and the other only on $\chi(\cos\theta)$. We may in fact write

$$F(\theta, t) = F(\theta, 0)\left[1 - \pi R^2 nt \sec\theta + \tfrac{1}{2}(\pi R^2 nt)^2 \sec^2\theta - \dots\quad\right] + X, \tag{39}$$

where

$$X = \pi R^2 nt\, \chi(\cos\theta) \sin\theta + \tfrac{1}{2}(\pi R^2 nt)^2 [\dots] + \dots \tag{40}$$

and so is a sum of terms depending only on $\chi(\cos\theta)$ and its differential coefficients. The coefficient of $F(\theta, 0)$ is readily summed, for on putting $\chi(\cos\theta) = 0$ our original differential equation (15) assumes the form

$$\frac{\partial}{\partial t} F(\theta, t) = -\pi R^2 n \sec\theta\, F(\theta, t),$$

of which the solution is

$$F(\theta, t) = F(\theta, 0)\, e^{-\pi R^2 nt \sec\theta}.$$

Thus the general value of $F(\theta, t)$, as given by equation (39), may be written in the form

$$F(\theta, t) = F(\theta, 0)\, e^{-\pi R^2 nt \sec\theta} + X. \tag{41}$$

We use this value for $F(\theta_0, t)$ in formula (28), and treat the two terms separately.

The first term to be integrated is

$$NF(\theta_0, 0)\, e^{-\pi R^2 nt \sec\theta_0}\, d\theta_0 \frac{1}{\pi\kappa t} e^{-\psi^2/\kappa t}\, d\psi\, e^{-\eta^2/\kappa t}\, d\eta, \tag{42}$$

and the total displacement θ is given as before by

$$\theta^2 = (\theta_0 + \psi)^2 + \eta^2.$$

For a divergent incident beam the integration can only be carried out by quadratures. If we simplify the problem by neglecting the divergence of the incident beam, then $F(\theta_0, 0)$ vanishes except when θ_0 is close to zero. Since we also have $\int F(\theta_0, 0)\, d\theta_0 = 1$, we may first integrate expression (42) with respect to θ_0, obtaining

$$Ne^{-\pi R^2 nt} \frac{1}{\pi\kappa t} e^{-(\psi^2 + \eta^2)/\kappa t}\, d\psi\, d\eta,$$

and then, since $\theta^2 = \psi^2 + \eta^2$, we may, by another integration, write this part of the law of distribution as

$$Ne^{-\pi R^2 nt}\, e^{-\theta^2/\kappa t} \frac{d\theta^2}{\kappa t}.$$

The second term to be integrated is treated just as before and the final result of the integration of this term is obtained from equation (37) on replacing $F(\theta, t)$ by X, or, what is the same thing, by taking in equation (35).

$$\Phi(\theta^2) = \frac{1}{\theta}\left[X + \tfrac{1}{4}\kappa t\frac{\partial^2 X}{\partial\theta^2} + \tfrac{1}{64}(\kappa t)^2\frac{\partial^4 X}{\partial\theta^4} + \ldots\right]. \tag{43}$$

With this value for $\Phi(\theta^2)$, the complete law of distribution is found to be

$$N\left[e^{-(\pi R^2 nt + \theta^2/\kappa t)}\frac{2}{\kappa t} + \Phi(\theta^2) - \tfrac{1}{2}\kappa t\frac{\partial\Phi(\theta^2)}{\partial\theta^2} + \tfrac{3}{8}(\kappa t)^2\frac{\partial^2\Phi(\theta^2)}{(\partial\theta^2)^2} - \ldots\right]\theta\, d\theta.$$

Integrating from 0 to θ and dividing by N, we obtain

$$\frac{I}{I_0} = e^{-\pi R^2 nt}(1 - e^{-\theta^2/\kappa t}) + \int_0^\theta \theta\Phi(\theta^2)\,d\theta - \left|\tfrac{1}{4}\kappa t\Phi(\theta^2) - \tfrac{3}{16}(\kappa t)^2\frac{\partial\Phi(\theta^2)}{\partial\theta^2} - \ldots\right|_0^\theta.$$

On substituting for $\Phi(\theta^2)$ from equation (43), and neglecting terms of order $\kappa^2 t^3$ and higher, this reduces to

$$\frac{I}{I_0} = e^{-\pi R^2 nt}(1 - e^{-\theta^2/\kappa t}) + \int X\,d\theta + \tfrac{1}{4}\kappa t\left(\frac{\partial X}{\partial\theta} - \frac{X}{\theta}\right). \tag{44}$$

As far only as first powers of t this becomes, on using the value for X given in equation (40),

$$\frac{I}{I_0} = 1 - e^{-\pi R^2 nt}\, e^{-\theta^2/\kappa t} - \pi R^2 nt\left[1 - \int_0^\theta \chi(\cos\theta)\sin\theta\, d\theta\right]. \tag{45}$$

On putting $\kappa t = 0$, this, of course, reduces to the formula (23), which was obtained before we took multiple-scattering into account. Thus the product of exponentials may be supposed to represent the effect of multiple-scattering to a first approximation.

Conclusion.

13. We have now examined the passage of a beam of particles which are subjected simultaneously to single-scattering of the kind first considered by Sir E. Rutherford, and to multiple-scattering of the type considered by Sir J. J. Thomson.

The question which, in view of recent experiments, is perhaps of greatest interest is that of whether there is, for small values of t, a range in which the intensity is of the type

$$I = I_0(1 - \epsilon t), \tag{46}$$

and, if so, that of the evaluation of ϵ. We have found that if multiple-scattering is entirely absent, there is such a range, and that ϵ will have the value calculated from the single-scattering theory on the supposition that the incident beam is strictly parallel and that no particle ever experiences more than one atomic encounter. This remains true even when the beam is not strictly parallel at incidence, but is scattered through a cone of semi-angle α, provided that α is so small that α^2 may be neglected. It also remains true when the particles are free, as in nature, to experience any number of atomic encounters in passing through the scattering substance.

When multiple-scattering is taken into account, there will also be a range in which the intensity is of the form (46), when the incident beam is not a strictly parallel beam but is diffused through a cone of a certain finite semi-angle α. But when α is small the range is very minute, and it disappears altogether when $\alpha = 0$ and the beam becomes strictly parallel. Throughout this range, when it exists, the value of ϵ is the same as though multiple scattering were entirely absent, and so may be calculated in the simplest way on the hypothesis of single-scattering.

When multiple-scattering is taken into account and the incident beam is strictly parallel, there is no longer a range in which the intensity is given by formula (46). The law of intensity for small values of t is now of the form (*cf.* equation (45))

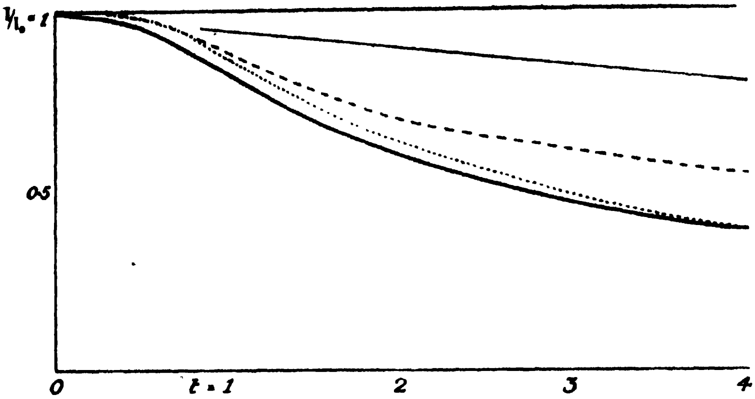
$$I = I_0(1-\epsilon t-e^{-\lambda t}\,e^{-\theta^2/\kappa t}). \tag{47}$$

The interpretation of this equation is obvious. The term ϵt represents the scattering produced by violent encounters; the factor $e^{-\lambda t}$ multiplying the last term is an extinction coefficient representing the rate at which the main beam is disintegrated owing to these violent encounters, while the other factor is the ordinary exponential of the theory of multiple-scattering, so that the product $e^{-\lambda t}e^{-\theta^2/\kappa t}$ represents the influence of multiple-scattering on a beam which is already somewhat disintegrated by violent atomic encounters. The two negative terms on the right-hand may thus be supposed to represent respectively the effects of violent encounters and of the "multiple-scattering" of a multitude of feeble encounters.

Of the three constants ϵ, λ, κ which enter into law (47), ϵ depends on θ, while λ and κ are independent of θ and depend only on the properties of the scattering substance. The constants ϵ and λ are not independent of one another, since λ is equal to the value assumed by ϵ when $\theta = 0$.

Fig. 3 shows the type of curve for I/I_0 predicted by equation (47). The thin straight line at the top is the graph of $1-\epsilon t$ with $\epsilon = \frac{1}{20}$; this would give the scattering for a certain value of θ and for certain material if multiple-scattering were neglected entirely. The dotted curve is the graph of $1-e^{-\theta^2/\kappa t}$ with $\theta^2/\kappa = 2$; this would give the scattering if single-scattering were neglected entirely. The chain curve is the graph of $1-e^{-\lambda t}\,e^{-\theta^2/\kappa t}$ with $\lambda = \frac{1}{10}$, and so gives the multiple-scattering after allowance has been made for the disintegration of the main beam by single-scattering. Finally, the thick curve (the lowest of all in the diagram) is the graph of I/I_0 as given by

Fig. 3



our equation (47). This is the type of curve which theory leads us to expect for small values of t.

We notice first how very closely this thick curve approximates to the dotted curve of pure multiple-scattering. The maximum difference (proportionally) between the ordinates occurs at about $t = 1{\cdot}5$, and is one of only 5 per cent. In our figure, the thick and dotted curves have been drawn for the same value of κ, but if we had obtained our thick curve as observational material, and had attempted, with κ entirely at our disposal, to draw a dotted (multiple-scattering) curve to fit it as closely as possible, we could have reduced this error to one of less than 3 per cent. It no longer appears surprising that the early experiments on the scattering of β-rays were thought to vindicate the theory of multiple-scattering.

We notice next that throughout a considerable range the thick curve can be fitted fairly closely by a straight line through the origin, so that if only a few points had been determined observationally on this curve it would be easy to mistake the curve for a straight line, and to believe, as later experimenters have done, that the experiments vindicated the theory of single-scattering. For instance, at $t \doteq 1, 3, 4$ the ratios of the total scattering to that predicated by the single-scattering hypothesis are respectively 3·42, 3·53, and 3·02. Any set of observations taken with thicknesses of foil in the neighbourhood of these values could be fitted on to a straight line with an average error of only about 6 per cent. If, however, this straight line were interpreted as a curve of single-scattering, the value of ϵ deduced from it would be in error by a factor of about 3·3.

At the outset I had hoped that this investigation would disclose some simple and practicable method of disentangling the effect of multiple-scattering from observational material, and so enabling us to deduce the value of ϵ, and hence the atomic law of force, with fair accuracy. So far from having disclosed any such method, our analysis suggests that it may be exceedingly difficult to detect the effects of single-scattering at all. We have seen that our full curve in fig. 3 can be fitted by a pure multiple-scattering curve to an accuracy of about 3 per cent., so that the problem of determining the effect of single scattering-reduces to that of measuring this 3 per cent. difference with accuracy.

It is, of course, clear that no general conclusion can legitimately be based on the single case which we have discussed and exhibited in fig. 3. There is, however, an underlying physical reason for the closeness of the two curves and the consequent difficulty of measurement, and this is of such a general nature as to render it futile to hope that other special cases would lead to substantially different results. The average β-particle, passing through the average experimental foil, experiences perhaps 10,000 encounters with atoms, of which one, or possibly two or three, will be the violent encounters of the theory of single-scattering. The theory of multiple-scattering is a purely statistical theory and is well able to deal with the 9,999 feeble encounters. In strictness it ought not to be asked to deal also with the one violent encounter; but it is a matter of common experience that we can frequently force the

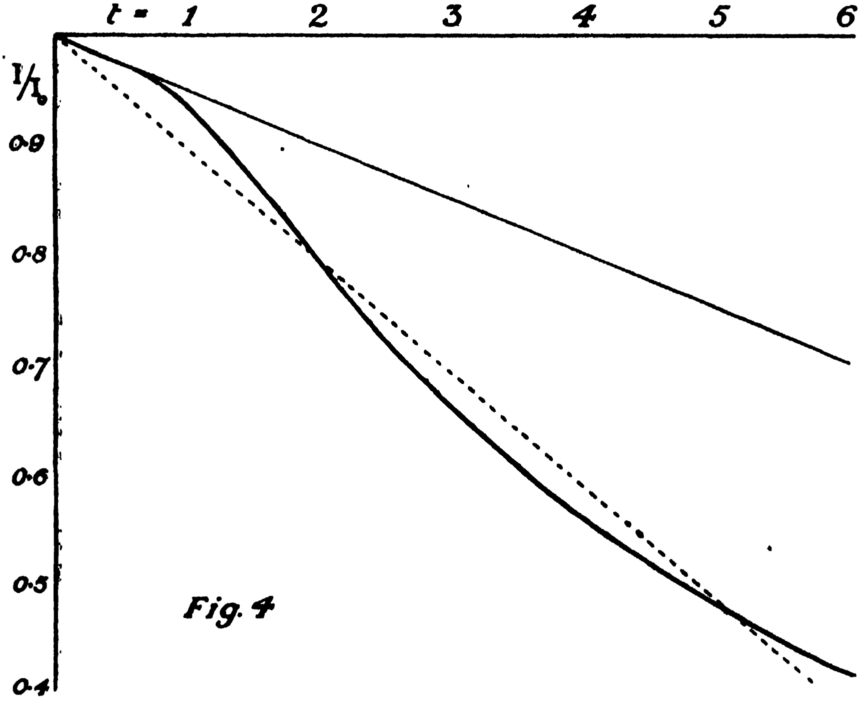


Fig. 4

statistical machine very considerably without breaking it. If the statistical machine is forced to deal with the whole 10,000 encounters, instead of with the 9,999 for which it is really suited, it merely regards the one big encounter as the equivalent of a large number of small encounters, and yields results which are not quite accurate, but are fairly so. The maximum difference of 2 or 3 per cent. between our full curve in fig. 3 and the adjusted dotted curve provides an instance of the smallness of the error. And it is on the detection and measurement of such small errors that a determination of the law of single-scattering must perforce rest.

In fig. 3 it will be noticed that the ratio of multiple to single-scattering (about 2·3) is somewhat greater than what probably occurs in practice. Fig. 4 represents the curve for the same amount of single-scattering ($\lambda = 1/10$, $\epsilon = 1/20$), but only about one-half of the previous amount of multiple-scattering ($\theta^2/\kappa = 4$ in place of $= 2$). The dotted curve is $I/I_0 = 1 - 2{\cdot}10\,\epsilon t$; its ordinate never differs from the true value of I/I_0 by as much as 8 per cent., and the average deviation is less than 3 per cent. Obviously it would be easy to mistake the thick curve for the dotted one and deduce a value of ϵ equal to 2·1 times the true value.

Spectrum of Active Nitrogen as Affected by Admixture of the Inert Gases, with a Note on the Origin of the Cyanogen Spectrum.

By Lord Rayleigh, F.R.S.

(Received October 27, 1922.)

(Plate 3.)

The visual spectrum of the afterglow accompanying the reversion of active nitrogen to ordinary nitrogen was investigated by Prof. A. Fowler and myself,* and it was shown that the remarkable green, yellow and red bands in the spectrum of this glow, originally observed by Lewis,† were in reality a selection of the first positive bands of nitrogen. We were only able to make this identification when we had examined the detailed structure of the bands on a fairly large scale spectrogram. Owing to the change of relative intensities the general aspect of the spectrum is so totally different from that of the positive column of a vacuum discharge, that it fails to suggest any connection. This is an unusual instance in spectroscopic experience. It is not often that the identity of a spectrum is totally disguised by changes in relative intensity.

---

* Fowler and Strutt, 'Roy. Soc. Proc.,' A, vol. 85, p. 377 (1911).

† 'Astrophysical Journal,' vol. 20, p. 49 (1904).

In 1912 I made experiments in search of similar afterglows in other gases than nitrogen. The results were, for the most part, negative, but in the case of helium a curious observation was made. The gas was contained in a vessel as shown in fig. 1, a powerful jar discharge passing between the electrodes. The gas diffused out of the region of discharge into the bulbous annexe, where it displayed the afterglow, if any. On introducing helium containing, as it happened, a little nitrogen, an orange-coloured afterglow was observed, which when examined with a small spectroscope appeared to be the same as that given by nitrogen. The redder colour was explained by a shift of relative intensity of the bands towards the red end of the spectrum. No measurement of the position of the bands was made. On absorbing the nitrogen with cooled charcoal, the afterglow became still redder, and finally disappeared. Pure helium gave no afterglow.

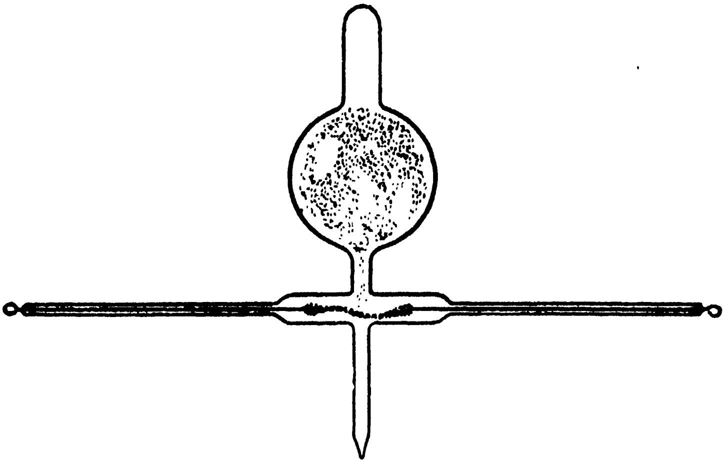


FIG. 1.

The matter was not pursued at that time; but, recently, trying to trace the origin of the red line in the aurora spectrum, I recalled these experiments and made a more detailed examination, which has revealed many curious and interesting facts. The same kind of vessel was used as before (fig. 1), and the spectrum of the glow in the bulb was photographed. The instrument used was one constructed for aurora work.* It had three 60° flint-glass prisms, and a Dallmeyer cinematograph lens working at F/1·9.

The bulb was well exhausted and heated, and various mixtures of nitrogen and helium were introduced, for photography of the afterglow spectrum.

---

* 'Roy. Soc. Proc.,' A, vol. 101, p. 115 (1922).

Exposures of about one hour were often necessary, using Ilford panchromatic plates.

It was often found rather troublesome to maintain the glow satisfactorily. A very small trace of carbon compounds causes the cyanogen spectrum to become conspicuous in the afterglow, and frequently after a few minutes run of the discharge this occurred. Whether the contamination came from the electrodes, the glass, or the stopcock used to shut off the vessel was not determined, but whatever the source may have been it made the experiment unsatisfactory.

[It has of recent years been doubted or denied that the spectrum in question is really due to cyanogen, as it is said to have been obtained in the absence of carbon.* On the other hand, Barrett† has traversed this conclusion, bringing forward evidence to show that in flame spectra this spectrum does not appear in the absence of carbon. I have not attempted to repeat or critically examine the experiments of Grotian and Runge, but the observations made with active nitrogen are very difficult to reconcile with their conclusion. A stream of nitrogen may be drawn through a discharge tube, by means of a motor-driven air pump, and under these conditions of constant renewal of the gas, the cyanogen spectrum in the afterglow may be reduced to a mere trace or got rid of entirely. On introducing any hydrocarbon into the stream of afterglowing nitrogen this spectrum invariably comes out with overwhelming intensity, and it may be seen to arise where the tributary stream of hydrocarbon enters, and *not* up-stream of this point. At the same time it may be shown by the Prussian-blue test that hydrocyanic acid is formed. These experiments have been described before,‡ but it is necessary to emphasize them in view of the subsequent controversy on the origin of the cyanogen bands.]

The spectrum became "clean" again after pumping out and renewing the gas, but it was troublesome to do this frequently. A remedy was found in admitting a carefully regulated trace of oxygen to oxidise the carbon compound. This caused the cyanogen spectrum to disappear completely. It was important not to add more oxygen than necessary, since if this is done the intensity of the glow is seriously affected.

A comparison of the photographs showed that the effect, on the afterglow spectrum, of adding helium to nitrogen was very peculiar. In order to be able to bring the various photographs into register, the soda line was put on to every plate as a comparison spectrum.§ In this way all possible doubt as to the identity of the bands, when comparing the photographs film to film, was avoided.

---

* See Grotian and Runge, 'Phys. Zeit.,' vol. 15, p. 545 (1914).

† 'Roy. Soc. Proc.,' A, vol. 98, p. 40 (1921).

‡ 'Roy. Soc. Proc.,' A, vol. 85, p. 288 (1911); also A, vol. 86, p. 112 (1911).

§ In some cases the lithium red line as well.

The spectra are reproduced in Plate 3, and the sodium line is indicated running through the series, but the result of a study of the intensities of the bands is more conveniently shown by a diagram, as in fig. 2, which shows the spectra in the same order as the plate. The diagram embodies all that can be seen on the original negative, which the reproduction of the latter can scarcely be expected to do.

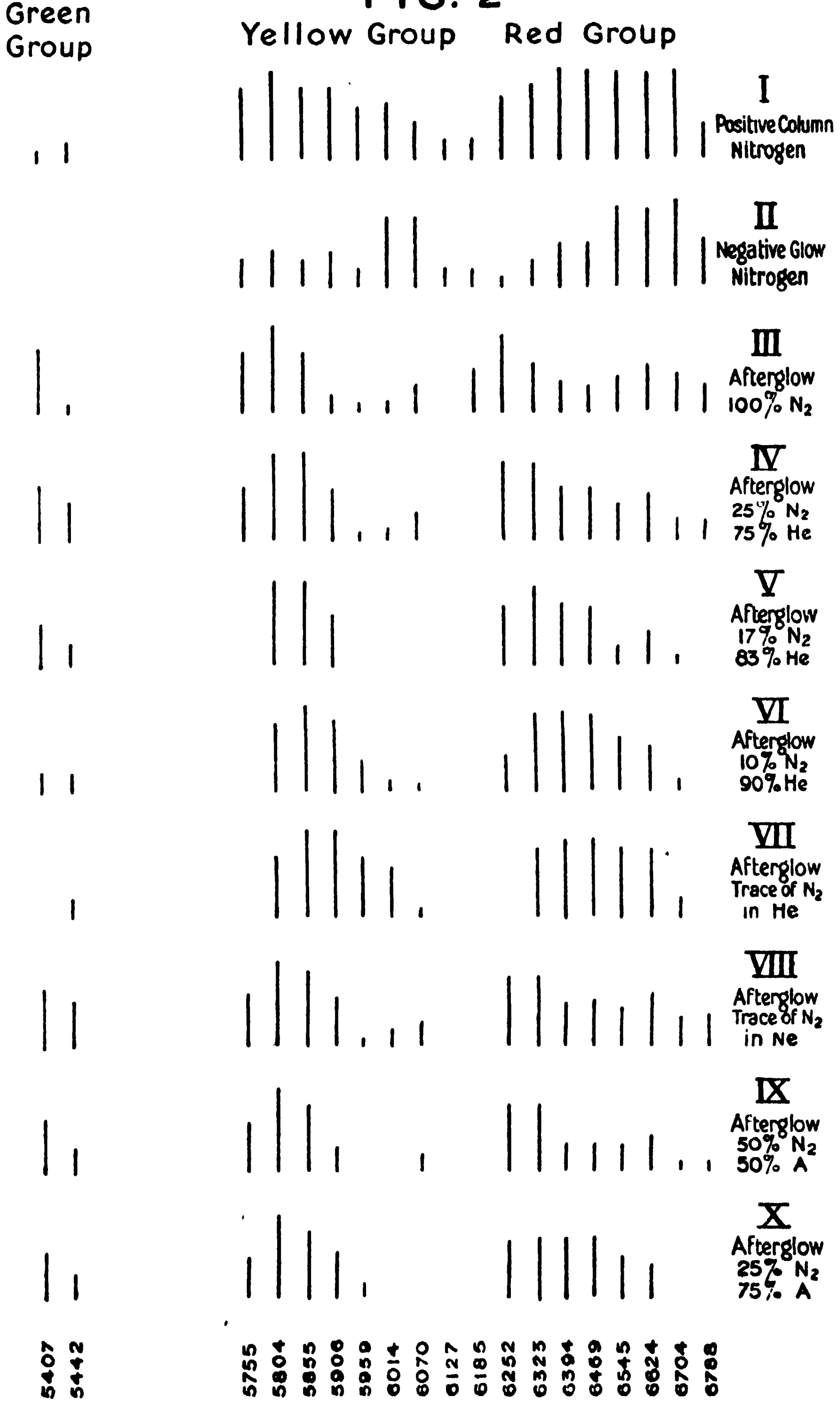
The length of exposure was such as to bring out the red group of bands to advantage, as it is here that the most interesting variations occur. The yellow bands, which are stronger, are somewhat over-exposed. It is probable that with still longer exposures even the faintest bands of the series might be made to show in every spectrum, and by taking a series of photographs with various exposures, material might be gained for exhibiting them all in order of intensity on a diagram. The limited range of gradation would prevent their being all shown on one photograph, at all events unless higher dispersion were used.

The diagram, fig. 2, shows the intensities on a scale of 10 by the vertical height of the lines. The horizontal scale is arbitrary, but roughly similar to that of the prismatic spectrum. In estimating the intensities of the bands in each spectrum, the strongest band or bands were marked 10, the next definitely weaker 9, and so on. It was found that ten steps were enough to include all the gradations that could be distinguished by inspection on the photographs available. If the weakest visible bands came higher than unit intensity, then the numbers were slightly readjusted, so as to duly represent the more striking contrasts to the best of the observer's judgment. An increase of unity on the intensity scale thus made probably represents rather a multiplication by some constant factor than the addition of a definite intensity.

The spectra obtained in this investigation were not on a sufficient scale for a study of the structure of the bands, but only of their general intensity. It may be that larger scale spectrograms, like those reproduced in 'Roy. Soc. Proc.,' A, vol. 85, p. 388, 1911, would show additional points of interest, and it would doubtless be possible to obtain them. This must be left for future investigation. The method of a flowing stream of gas, formerly used, gives a brighter afterglow, but is not so well adapted when rare and expensive gases are used.

It will be convenient first to draw attention to the ordinary first positive spectrum of nitrogen as obtained in the capillary of a spectrum tube, in which the majority of the bands shown are of a tolerably uniform intensity

FIG. 2


Green Group
Yellow Group
Red Group
I
Positive Column Nitrogen
II
Negative Glow Nitrogen
III
Afterglow 100% N_2
IV
Afterglow 25% N_2 75% He
V
Afterglow 17% N_2 83% He
VI
Afterglow 10% N_2 90% He
VII
Afterglow Trace of N_2 in He
VIII
Afterglow Trace of N_2 in Ne
IX
Afterglow 50% N_2 50% A
X
Afterglow 25% N_2 75% A
5407
5442
5755
5804
5855
5906
5959
6014
6070
6127
6185
6252
6323
6394
6469
6545
6624
6704
6788

(spectrum I). The next spectrum (II) shows the same bands as they appear in the negative glow of a discharge tube. The intensity distribution is now strikingly different, especially in the development of λ 6014 and λ 6070.

We come now to the main subject of the observations. No. III is the spectrum of the afterglow when nitrogen is introduced into the special discharge tube (fig. 1). The bands are conveniently regarded as belonging to a green, yellow, and red group, as indicated at the top of fig. 2. Visually, in a small instrument this distinction is even more apparent than on the photographs, as only the stronger bands of each group are readily seen; the weaker bands, linking the yellow and red group, only come out photographically with long exposure. Moreover, the green group is visually even stronger than the red one.

It is necessary to add an excess of helium in order to produce any perceptible effect on the intensity distribution. No. IV shows the afterglow spectrum with a mixture of helium and nitrogen containing 75 per cent. of the former by volume. Speaking generally, the effect of adding helium may be said to be a shift of the maximum of intensity in each of the three groups towards the red end of the spectrum.

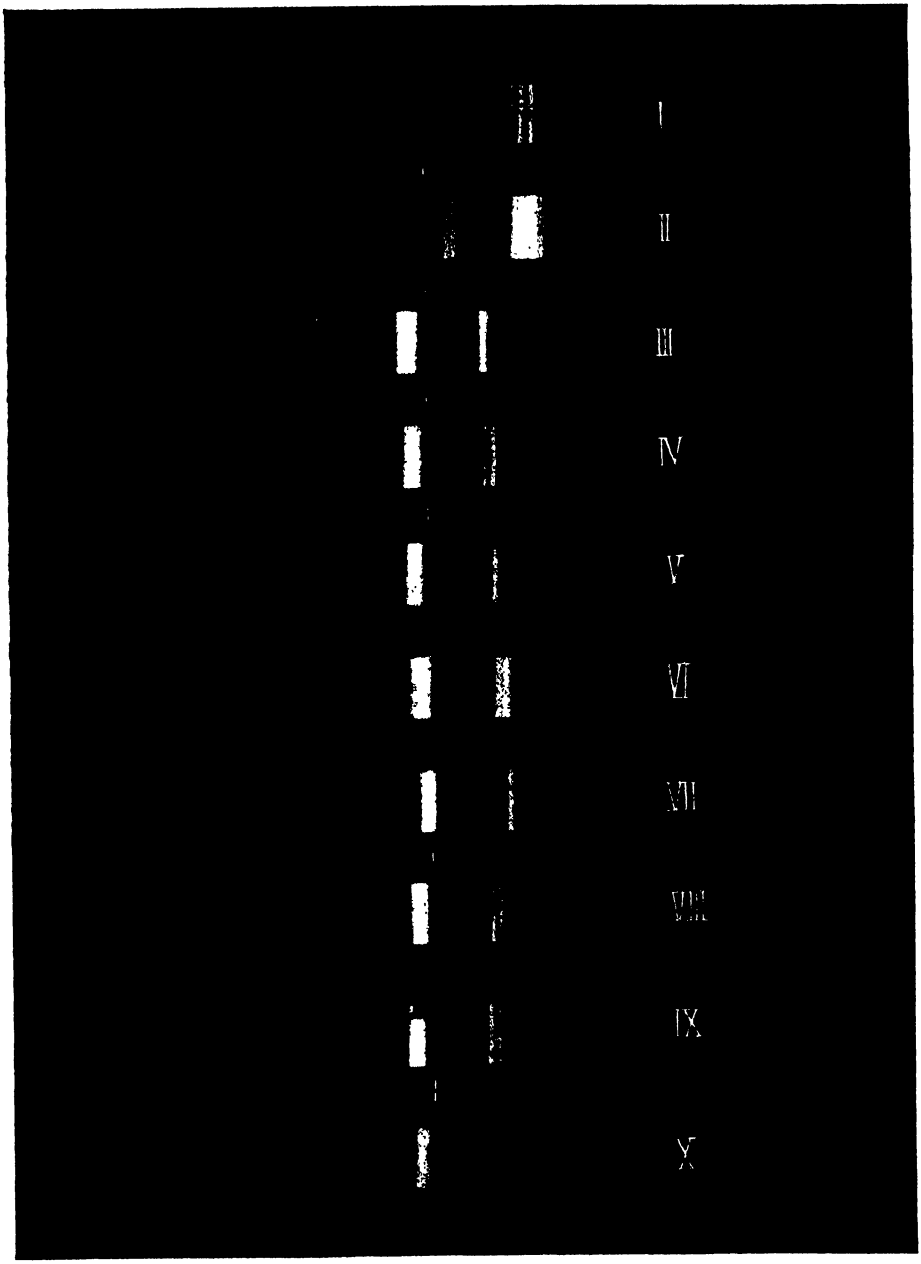
No. V is the next stage (83 per cent. of helium by volume). The shift towards the red has proceeded further.* The green mercury line 5461 is present on the photographs V, VI, VII, but is not indicated in fig. 2. It is due to accidental presence of mercury, which is stimulated by the active nitrogen.

Spectrum VI, 90 per cent. of helium by volume, shows a further movement of the maximum in each group towards the red. Moreover, there is a marked gain of intensity in the red group and loss of intensity in the green group, relative to the yellow group. This corresponds with the general reddening of the afterglow, originally noticed by the unaided eye.

Spectrum VII.—In this case the helium used had been supposed to be fairly free of nitrogen, but it evidently contained a little—exactly how much was not determined, but perhaps something of the order of 3 per cent. It is remarkable how bright the afterglow remains under these circumstances,

---

* This spectrum is of interest in connection with the aurora. Dr. V. M. Slipher, of the Lowell Observatory, U.S.A., has communicated to me an approximate photographic determination of the wave-length of the red aurora line as λ 6320, in close agreement with the strongest red band in the above spectrum. In this connection it is to be remarked that an important part of the aurora spectrum is a spectrum of nitrogen, and that we have reason to believe that there is an excess of helium in the upper atmosphere. Putting these considerations together, I had hopes that the origin of the red aurora line had been traced. A small-scale spectrum, obtained for me at Shetland by Mr. J. Crichton, shows the red aurora line, but its general aspect did not encourage the suggested identification. It is hoped to obtain a better photograph of the red aurora line this winter which may be more decisive.


I
II
III
IV
V
VI
VII
VIII
IX
X

but it had now acquired a decidedly red colour. There are now five fairly strong bands in the red group.

VIII is the spectrum of the afterglow in a sample of gas rich in neon, which I owe to the kindness of the Société Anonyme d'Air Liquide, of Paris. The gas was used as received, after absorption of any oxygen it might contain. The percentage of nitrogen is not known, but was almost certainly a good deal less than the neon. Within the limits of experimental erro the spectrum is practically identical with No. IV, obtained with a 75 per cent. helium mixture.

Spectrum IX, with 50 per cent. of argon in the nitrogen, is nearly the same as VIII (neon) and IV (helium).

Spectrum X, with 75 per cent. argon, is of interest because it shows a distribution of intensity which cannot be obtained with any helium mixture, whatever the composition. When argon is added no general reddening of the glow is observable. It was found that mixtures still richer in argon did not give a glow bright enough to be easily investigated.

Summary and Conclusion.

It has been shown in earlier investigations that the usual spectrum of the nitrogen afterglow consists of a selection of the first positive nitrogen band spectrum. It is now shown that this selection can be modified by large admixture of the inert gases. There are three groups of bands, green, yellow and red, and the effect of adding an inert gas is to shift the maximum of intensity in each group towards the red. Addition of helium in particular makes the red group, as a whole, more intense at the expense of the others.

The initial effect of adding helium, neon or argon is the same, but an excess of argon produces effects different in detail from those produced by an excess of helium.

I have not succeeded in determining whether these effects depend on something which occurs in the electric discharge, or on something which happens later, in the process of the afterglow.* If the latter is the explanation, the effects might be attributable to something of the nature of a selective damping of certain periodicities when collisions occur; but in the present state of knowledge the time is hardly ripe for speculating further. It is believed that the facts obtained cannot fail to be of considerable significance.

In addition to the main subject of the paper, evidence is recapitulated which shows the influence of carbon in developing the "cyanogen" bands.

---

* To do this would require more powerful air-pumps than I have now at command.

Directional Observations of Atmospheric Disturbances, 1920–21.

By R. A. WATSON WATT, B.Sc., F.Inst.P.

(Communicated by Sir Henry Jackson, F.R.S. Received June 27, 1922.)

(Published by permission of the Radio Research Board.)

Introductory.

The paper analyses 12 months' observations of the intensity and apparent direction of arrival of atmospheric disturbances, carried out for the Radio Research Board at a station in S.E. England. Observations were made at 7 h., 13 h., and 16 h. daily, the average number of such disturbances heard being 80 per minute. The apparent direction of arrival is well marked, with unexpectedly simple diurnal and seasonal variations, following a general law of counter-clockwise swing in direction accompanying increase in solar altitude. The mean direction of greatest disturbance for the year is 153°, with a diurnal variation of 60° range, a seasonal of 100°.

The numbers heard per unit-time have also comparatively simple diurnal and seasonal variations, March being the quietest, June the most disturbed month, with a disturbance twice that of March. Fourier coefficients for the monthly mean and modal values are tabulated. Close correlation between apparent direction of arrival and solar altitude is found, no local topographical effect is traced, and no quantitative relation with terrestrial magnetism, atmospheric electricity, or any of the surface meteorological elements (except temperature) has yet been established. Special observations during a solar eclipse are reported.

I. In the course of a Meteorological Office investigation into the location of thunderstorms by directional radio-telegraphic observation, it soon became clear that the extent to which thunderstorm detection by such means could be made practical, depended on the results of a general study of the naturally occurring electro-magnetic waves, variously called "atmospherics," "strays," "X's," "sturbs," and "static." The first steps in such an investigation were accordingly made in 1916, and continued as the exigencies of other work permitted.

The formation of the Radio Research Board in 1920, and the appointment by that board of a Sub-Committee on Atmospherics, resulted in the transfer of the Meteorological Office Radio Station at Aldershot to the administration of the Radio Board for the prosecution, in the interests of scientific and technical radio-telegraphy, of the more general investigation referred to.

Geophysical investigations, in general, share a common difficulty in the vast and uncontrollable extent of the phenomena dealt with; the problem of determining the origin and nature of atmospherics adds to this the instrumental limitations of high-frequency alternating current measurement, which, in this case, must deal with highly damped impulses of very short duration.

There were, however, certain characteristics of the general stream of atmospherics which could be investigated without much delay, and the present paper deals with a series of observations, covering a whole year, of the apparent mean direction of arrival of atmospherics, and their numerical frequency.

The observing station at Aldershot has co-ordinates lat. 51° 16′ N., long. 0° 46′ W., height above M.S.L. 85 metres, and is situate on moderately open heath, with but few trees.

The apparatus used consisted of an indoor frame antenna, approximately 2 metres square, swung on a vertical axis, and tuned to a wave-length of 20,000 metres. This relatively high wave-length was chosen to avoid interference from ordinary radio-telegraphic traffic, and to take advantage of the fact that atmospherics are more numerous and more intense at long wave-lengths. A considerable number of observations have been made on shorter wave-lengths, in a great majority of cases no difference in apparent direction was found on changing from 20,000 to 2,000 metres. The intensity of disturbance was less on the shorter waves, but the general features of the distribution were unchanged.

The atmospherics were amplified and detected in a multi-stage triode amplifier, specially chosen for its freedom from internally generated noises, which, in many amplifiers, closely simulate atmospherics.

Observations were made, in general, three times daily, at 7 h., 13 h., and 16 h., G.M.T., but certain week-end observations were missed. The 7 h. series, however, is uninterrupted. The method of observation was to set the frame with its plane in the meridian, count all atmospherics heard in 15 seconds, then set the frame at 30°, 60°, 90°, 120°, and 150° to the meridian, and make similar counts, finally counting with the frame 180° from its original setting. These seven observations constitute the "primary count." The frame was then set at 10° steps round the setting at which fewest atmospherics were heard in the primary count, then at 5° steps, until the position of minimum disturbance was located to 5°, or, when possible, to 2½°, with the usual radio-goniometric ambiguity of 180°. A second count was then made with the frame set at 90° from the position of minimum disturbance, and, to eliminate instrumental error, at 180° and 270° from minimum.

The 15-second time-unit was adopted as giving numbers of convenient magnitude, numbers per minute being inconveniently large.

Any separately distinguishable sound of the "atmospheric" type, *i.e.*, anything recognisable aurally as a peak in a curve of audibility, whether apparently discrete or connected to another such peak by an indeterminate sound, was counted as one atmospheric. Apart from this convention, and an occasional distinction between "grinders" and "clicks," no effort was made to separate the atmospherics into types or intensity classes. Considerable experience has shown the impossibility, with existing aural methods, of obtaining satisfactory directional observations when such separation is attempted. The observations were made exclusively by two workers, who made joint observations at intervals to maintain uniformity in standards of estimation.

II. The observations discussed, extending from November 1, 1920, to October 31, 1921, number in all 941, of which 365 were at 7 h., 310 at 13 h., and 266 at 16 h. Of the 7 h. observations, 101, or 28 per cent., fell before sunrise; of the 16 h.—17 h. observations (which were more variable in time because of interference from local transmission), 26, or 10 per cent., fell after sunset.

The distribution of observations is set out in detail in Table 0.

The total number of atmospherics individually enumerated in primary counts was 133,023, an average of 80 per minute, or 4×10^7 per annum. Apart from neglect of minor disturbances, these figures are an underestimate, because the observer's counting powers are found to break down at about 140 per minute, so that counts of rates exceeding this not infrequent value were entered as "continuous" and treated statistically as equalling 140 per minute.

Attention should again be drawn to the fact that the determinations have an ambiguity of 180°, and that, except where otherwise stated, the directions of minimum atmospheric disturbance will be taken, conventionally, as lying between 0° and 180° E. of N. In cases where a resolution of the ambiguity is necessary for statistical purposes, the sense is chosen to give the nearest approach to symmetry about the most frequent value, *i.e.*, all directions are assigned to a range 90° on either side of this value.

III. The frequency curves, figs. 1 and 2, show the number of occasions on which a given azimuth from true north was located, by the methods described, as the least disturbed direction, in groups containing all observations at each of the three fixed hours, and all in the three seasons, winter, summer, and equinox respectively. The corresponding assumed direction of maximum disturbance, 90° clockwise from the minimum, is shown on a subsidiary scale. The minimum, being much more sharply defined, is the actual observed quantity, but occasions on which the direction of maximum disturbance is observed to differ from the minimum by an angle other than 90° are very rare. The frequency Table (Table I) shows the direction, N_F, forming the peak of smoothed frequency curves for the three fixed hours

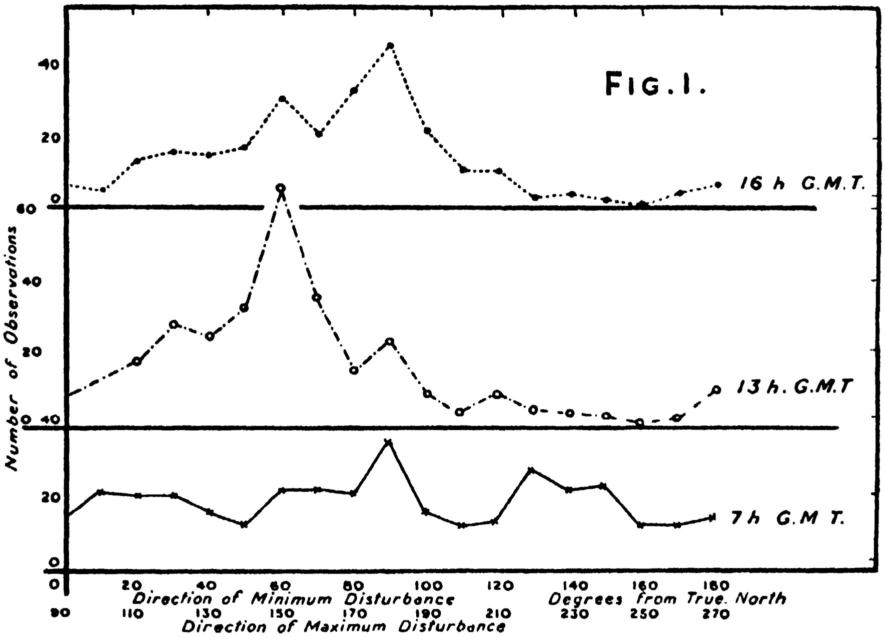


FIG. 1.

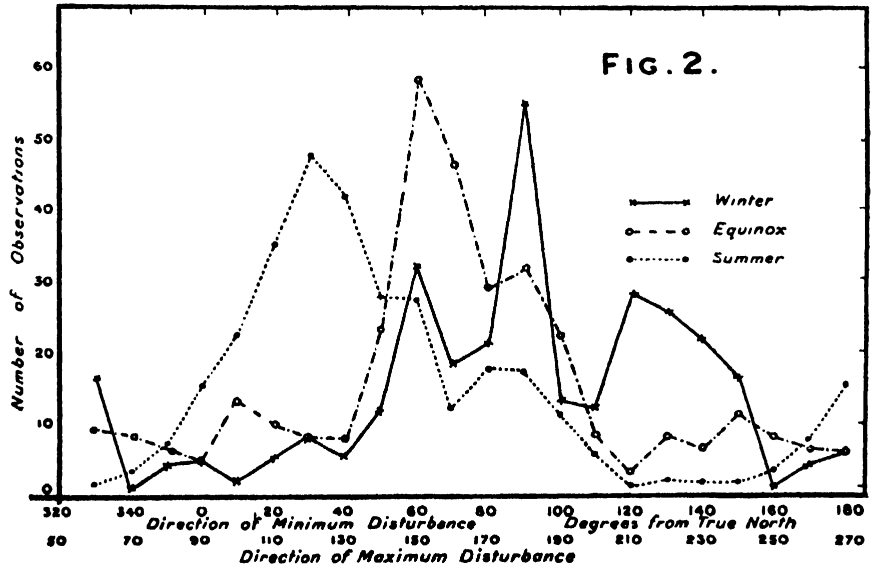


FIG. 2.

Table No. ...	0.				I.				II.				III.				IV.			
	Number of observations.				Most frequent minimum direction N_F.				Mean direction of minimum N_M.				Mean No. per ¼ min. at instantaneous minimum.				Mean No. per ¼ min. at instantaneous maximum.			
Time (G.M.T.)...	7.	13.	16.	Total.	7.	13.	16.	Total.	7.	13.	16.	Total.	7.	13.	16.	Total.	7.	13.	16.	Total.
November.........	30	26	22	78	60	70	90	70	140	57	80	60	17·2	19·2	18·6	18·3	24·0	26·9	25·7	25·5
December.........	31	26	21	78	142	58	132	130	140	70	110	125	17·3	16·0	16·7	16·7	25·3	23·1	23·7	24·0
January	31	26	21	78	90	90	90	90	115	100	100	100	16·3	15·9	15·7	16·0	23·7	22·7	21·6	22·7
February	28	23	20	71	134	61	65	68	110	65	80	80	15·0	13·2	12·7	13·6	20·6	18·2	17·7	18·8
March	31	27	21	79	134	66	81	67	100	70	80	70	12·6	10·4	10·9	11·3	17·0	15·1	15·3	15·8
April..............	30	25	20	75	120	58	65	62	85	70	70	70	13·2	11·5	12·9	12·5	17·2	16·3	18·1	17·2
May	31	26	23	80	67	52	47	54	85	15	45	50	13·9	17·2	17·4	16·2	17·7	23·9	24·6	22·1
June	30	26	23	79	13	20	34	23	20	10	20	20	16·5	22·2	22·5	20·4	21·4	29·7	29·2	26·8
July	31	25	21	77	34	24	21	23	40	20	30	30	15·3	20·3	21·1	18·9	19·8	27·2	27·3	23·8
August	31	25	23	79	30	37	39	35	43	37	37	36	16·3	18·5	17·4	17·4	20·8	24·2	22·4	22·5
September.........	30	28	26	84	10	59	84	60	24	50	86	56	18·3	18·9	20·2	19·1	23·1	24·1	25·5	24·2
October............	31	27	25	83	154	65	87	67	162	64	92	73	17·3	17·8	19·3	18·1	22·3	23·7	24·7	23·6
12 months......	365	310	266	941	89	61	87	62	128	52	68	61	15·8	16·8	17·4	16·6	21·0	23·0	23·3	22·3
					Figures derived from frequency curves for each period.				Figures derived from distribution curves for each period.				Total $= \frac{7+13+16}{3}$, to avoid preponderance of 7 h. values.				Total $= \frac{7+13+16}{3}$.			

Table No. ...	V.				VI.				VII.				VIIA.				VIII.			
	Mean No. per ¼ min. at N_M.				Mean No. per ¼ min. at $X_M = N_M + 90°$.				Mean instantaneous directivity ratio.				Mean directivity ratio.				Audibilities.			
Time (G.M.T.)...	7.	13.	16.	Total.	7.	13.	16.	Total.	7.	13.	16.	Total.	7.	13.	16.	Total.	7.	13.	16.	Total.
November	20·2	21·4	21·6	22·0	22·8	25·7	24·4	24·9	28·3	28·6	27·6	28·2	11·4	16·7	11·5	11·6				
December	20·3	19·3	20·4	20·7	23·6	21·5	22·8	22·0	31·6	30·7	29·6	30·6	14·0	10·2	10·5	5·9				
January	18·2	17·0	16 6	17·3	23·0	22·0	21·0	22·0	31·2	29·9	27·3	29·5	20·8	22·7	20·9	21·4				
February	16·6	14·3	14·6	15·5	19·7	17·6	17·2	17·8	27·2	27·4	28·2	27·6	15·7	18·8	15·1	12·9	1·06	0·72	0·51	0·76
March	14·4	11·1	12·0	12·7	16·4	15·0	15·0	15·5	23·8	31·1	28·8	28·6	12·2	15·6	20·0	18·1	0·47	0·38	0·26	0·37
April	14·7	12·3	13·6	13·4	16·7	16·1	18·0	17·0	23·2	29·4	28·7	27·1	12·0	23·6	24·4	21·2	0·20	0·27	0·31	0·26
May	15·2	18·5	19·1	18·0	17·2	23·8	24·2	21·5	21·5	28·1	29·3	26·3	11·6	22·2	21·1	19·4	0·30	0·76	0·92	0·66
June	18·6	23·9	24·8	22·7	20·6	28·4	28·5	25·8	22·9	25·3	23·0	23·7	9·7	15·8	13·0	12·0	0·26	0·41	0·47	0·38
July	17·2	21·6	23·0	20·8	19·0	26·8	26·4	23·9	22·7	23·4	22·7	23·6	9·5	19·4	12·9	13·0	0·25	0·49	0·50	0·41
August	17·9	20·0	20·5	19·4	20·2	23·8	23·8	22·6	21·9	23·4	22·8	22·7	11·4	16·0	14·7	14·1	0·31	0·36	0·43	0·37
September.........	19·5	20·8	21·1	21·1	25·5	24·0	22·9	23·5	20·6	21·4	20·5	20·8	23·5	13·3	7·8	10·2	0·36	0·52	0·47	0·45
October	18·2	18·4	20·3	20·5	21·9	23·6	22·4	21·9	22·2	24·8	21·6	22·9	16·9	22·0	9·4	6·7	0·76	0·84	0·81	0·80
12 months......	18·1	13·6	20·0	19·3	19·6	22·0	22·4	20·7	22·1	24·5	23·2	23·2	15·4	15·4	10·7	6·8				
	From distribution curves.				From distribution curves.				From tabulated directivity ratios for each observation. Total = $\frac{7+13+16}{3}$.				Mean directivity ratio = $\left(\frac{\text{No. at } X_M - \text{No. at } N_M}{\text{No. at } X_M}\right) \times 100$.							

for each month and for the year, N_F being thus the most frequent direction of minimum atmospheric disturbance for the period. Parts of the frequency curves are repeated, where necessary, to give an unbroken range of $N_F \pm 90°$. Throughout the paper the words mode and modal are used when referring to most frequent as distinguished from mean values.

The curves are all seen to be distinctly skew, with, in general, a steeper gradient on the side of counter-clockwise deviation. This appears to be associated with a secondary peak, relatively prominent in winter and spring, and in 7 h. curves, in the region of 130°, while the main peaks are in the 30°—90° region. The skewness is found to be less in the case of the four-month seasons than in that of the three-month seasons. In fact, the general symmetry of fig. 2, for the four-month seasons, is the first indication of a marked solar control of the phenomena under consideration.

IV. *Mean Direction Tables.*—The primary counts for the various periods have been meaned, and the mean directions of minimum disturbance are set out in Table II as N_M. This direction, N_M, is the trough of a period curve in which the mean number of atmospherics heard with the frame set at azimuth θ is plotted against θ, which has successive values of 0°, 30°, 60°, ..., 180°. Such a curve, formed by the superposition of curves from individual primary counts, is flatter than the instantaneous curve, and becomes flatter the longer the period, but the trough can be located within 5°, and the agreement between N_F and N_M is, on the whole, close.

V. *Curves of Monthly Values.*—The data of Tables I and II are further summarised by the curves of figs. 3 to 5. In all these curves the abscissæ are months; the ordinates are: in fig. 3 the most frequent direction of minimum disturbance (N_F) for each of the three fixed hours; in fig. 4 the corresponding mean direction, N_M; and in fig. 5 the arithmetical mean of the fixed hour ordinates are plotted to bring together two curves representing, as nearly as possible, N_F and N_M freed from diurnal variation.

VI. *The Instantaneous Directivity Ratio.*—The improvement, in freedom from interference by atmospherics, obtainable by setting a frame antenna to the instantaneous direction of minimum disturbance may be measured by the instantaneous directivity ratio, which, as a percentage, is expressed by

$$\frac{\text{No. per unit time at maximum} - \text{No. at minimum}}{\text{No. at maximum}} \times 100.$$

It expresses the percentage of all atmospherics received which are directed, on the crude assumption that the distribution consists of a number, equal to the number at minimum, coming in uniformly from all azimuths, plus a number, equal to maximum minus minimum number,

FIG. 3.

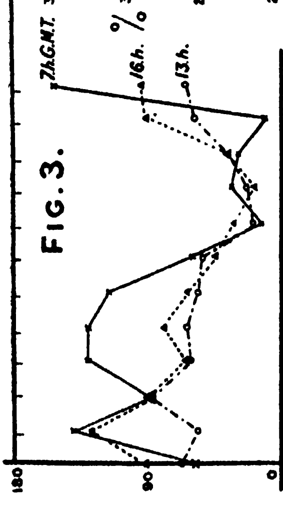


FIG. 4.

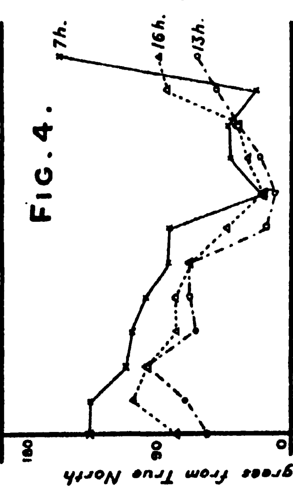


FIG. 5.

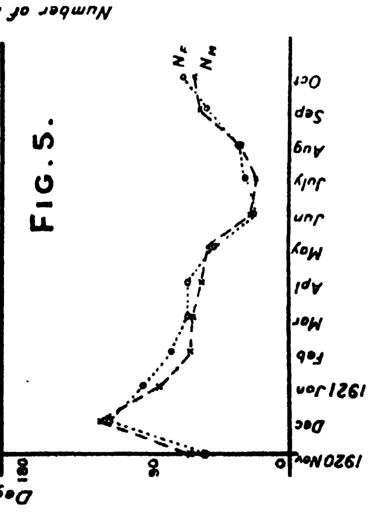


FIG. 6.

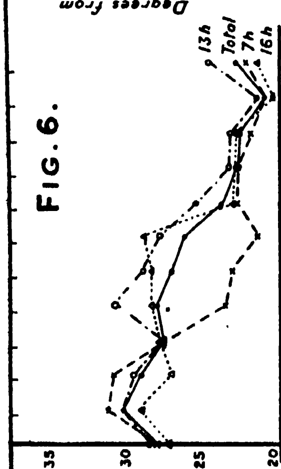


FIG. 7.

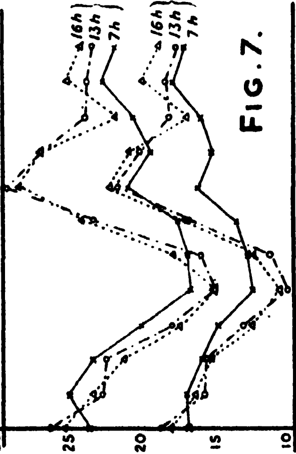


FIG. 8.

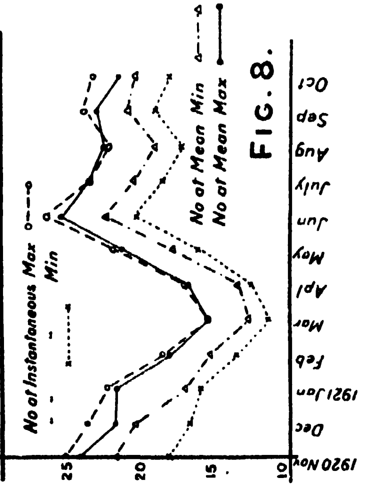
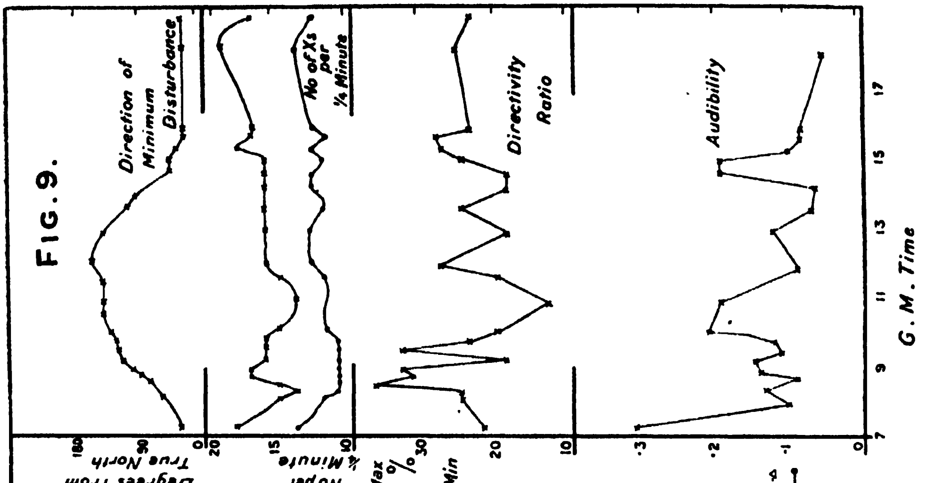


FIG. 9.



coming solely from the direction of maximum. This ratio is the only one which suggests itself as useful in the present state of direction-finding on atmospherics. It neglects the complications due to the vertical receptivity of the frame, and to the probably complex distribution in azimuth.

Table VII shows the monthly and annual mean value of this ratio; and monthly mean values are plotted in fig. 6.

A mean directivity ratio may be similarly defined by

$$\frac{\text{Mean No. at maximum} - \text{Mean No. at minimum}}{\text{Mean No. at maximum}} \times 100.$$

Using monthly means, this ratio is of the order of 15 per cent., but its significance in practice is slight.

VII. *Intensity Curves and Tables.*—Tables III and IV show the monthly mean values of the numbers of atmospherics heard per unit-time at the instantaneous positions of minimum and maximum disturbance respectively. Tables V and VI show similar mean values referred to the corresponding mean positions, N_M and X_M ($= N_M + 90°$), *i.e.*, they represent the minimum and maximum ordinates of the monthly distribution curves plotted from the primary counts, these ordinates occurring at abscissæ N_M and X_M respectively. The type of variation is sufficiently represented by fig. 7, while the curves of fig. 8 show the relatively smooth and similar run of the seasonal variation when freed from diurnal effects.

VIII. *Fourier Coefficients.*—It was not to be anticipated, *a priori*, that a short period of observation, such as one year, on an apparently random and certainly complex geophysical phenomenon, would yield results easy of interpretation. It is, then, encouraging to note the relative smoothness of the curves of annual variation, both in direction and intensity.

Figs. 5 and 8 are of a smoothness which one might almost class as astronomical, and are direct evidence of a fundamentally simple type of control. Fig. 6 is somewhat less promising, but all the variation curves are such as to be adequately representable by a very few terms in a Fourier series.

The variation shown by the twelve monthly ordinates has accordingly been analysed, in each case, into the series

$$F = C_0 + C_1 \sin(\phi + x_1) + C_2 \sin 2(\phi + x_2) + \ldots$$
$$+ C_6 \sin 6(\phi + x_6) \ldots \text{ when } x_6 \text{ is } 15°.$$

The coefficients, C_n, and phase angles, X_n, are set out in Table IX, the zero date being June 21 in all cases. The approximate date of maximum of each term is also tabulated for reference. Their main features are noted in the succeeding sections.

IX. *The Diurnal and Seasonal Variation of Direction.*—The prominent features of the directional variations are the extreme variability of 7 h. data as compared with the other hours, and the smoothness of variation in these other cases. Freed from diurnal effects, the variation takes the form of a steady swing of the most disturbed direction from 215° (S.W. by S.) in December, through 160° in March to 110° in June. The range of diurnal variation may amount to 60° or 80°.

It is convenient to adduce here the evidence which justifies the assumption already made, and repeated in the preceding paragraph, of a sense, for the ambiguous direction of arrival, corresponding to arrival from lower latitudes. On general grounds, it would be difficult to find a tenable theory of the origin of atmospherics which did not lead to a direction of travel from lower to higher latitudes, but experimental evidence is not lacking.

The writer made, in 1917, a series of observations in which his unambiguous direction-finder was used. In more than 80 per cent. of the observations, lasting over several months, the average direction of arrival of atmospherics was from the southward.

Further, a considerable mass of data from the Admiralty coastal direction finding stations (now being prepared for publication) gives an almost invariable southward convergence of the bearings of atmospherics simultaneously observed at several stations, thus indicating a travel from lower to higher latitudes.

Reverting to the variations in apparent direction, the Fourier series presents the following features:—

The direction of maximum disturbance for the whôle year lies between 152° and 154°, according to whether the modal or the mean value is considered. The non-periodic term of the 7 h. series is 172° or 179°, of the 13 h. series 145° or 142°, and of the 16 h. groups 160° or 164°.

This term (C_0) has accordingly a diurnal variation such that the apparent direction of arrival of atmospherics swings counter-clockwise with increasing solar altitude, and *vice versâ.* The twelve-month term (whose amplitude, C_1, is of the order of half the constant term, and always more than twice that of the six-month term, so that the variation of direction is quite well represented by the constant and the twelve-month term) has its clockwise maximum some week or two after the winter solstice, and has amplitudes of the order of 50° at 7 h., 25° at 13 h., and 36° at 16 h. This annual variation is therefore in the same general sense as that of the diurnal variation of C_0; the direction of arrival swings counter-clockwise, as regards both day and season, with increasing solar altitude. The diurnal variation of C_1 would fit with the fact that the 7 h. readings were frequently before sunrise, the

16 h. less frequently after sunset, the 13 h. in full daylight. The amplitude, C_1, therefore, is in some measure determined by the annual variation of conditions of light and dark.

Throughout the analysis, modal and mean values have been treated separately, in the hope of throwing light on the skewness of the frequency curves. It is notable that the effect of this skewness is relatively slight at 16 h., and that there is some symptom of its affecting the odd harmonics less than the even.

X. *The Diurnal and Seasonal Variations of Intensity.*—The most pronounced feature of the diurnal variation is the reversal of the law which applied to direction; the 7 h. observations, which were relatively variable in direction, are relatively invariable in intensity. From December to April the disturbance at 7 h. is greater than that at 13 h. or 16 h.; from May to November the 7 h. disturbance is much less than that later in the day. The 7 h. curve has a maximum in December and January, in which months all 7 h. observations occurred before sunrise, falls to a minimum in March–April, reaches a secondary maximum in June, and rises to its main maximum in September.

The 13 h. and 16 h. curves have a minimum along with the 7 h. curves in March, then rise to a well-marked maximum in June.

The range of variation is large in summer, the numbers increasing by some 40 to 50 per cent. between 7 h. and 13 h.; at other seasons the range of the reversed variation is comparatively small.

The elimination of the diurnal variation, by taking the arithmetic mean of the fixed hour means, gives curves (fig. 8) of a regularity inferior only to that of the directional curves, and still much better than was, *a priori*, anticipated. These curves show a marked minimum in March and a maximum in June. The second equinoctial trough, which might have been looked for, is absent, and is replaced by a distinct secondary maximum. This may be fortuitous, but the meteorological vagaries of September have been a subject of discussion in recent years.

The disturbance in June is nearly double that of March. This somewhat curious rapid change of conditions between spring and summer, with a relatively slow recovery throughout autumn and winter, is paralleled by directional data of atmospherics on short wave-lengths during four years referred to in paragraph IX.

The Fourier coefficients (see Table IX) reflect these features, the 7-h. data being fairly well represented by a twelve-month term only, and quite adequately by the addition of a six-month term; the 13-h. and 16-h. disturbance by twelve- and six-month terms, with an unimportant four-month term. The six- and four-month terms are relatively prominent at 13 h.

The diurnal variation of the non-periodic term C_0, and of the twelve-month amplitude C_1, is in the direction of an increase in value from 7 h. to 13 h. and from 13 h. to 16 h.

The twelve-month term has its maximum near the autumnal equinox, the others have maxima near the summer solstice.

The six-month term is more prominent at the direction of maximum disturbance than at the minimum.

XI. *Instantaneous Directivity Ratio.*—The close agreements between the figures for intensity at instantaneous direction of minimum disturbance, N_1, and at the direction of maximum, $X_1 (= N_1 + 90°)$, both as to relative amplitudes and phase angles, indicate, if such indication were necessary, the fact that we are dealing with a single distribution, and that the assumed sharply directed stream of atmospherics is not subject to a different set of laws from the assumed random background. This would appear to transfer the interest of the directivity ratio from the source or origin of the atmospherics to the properties of the medium of propagation and the means of reception, since it makes it probable that the variation in sharpness of definition of the assumed stream, of which the directivity ratio is a measure, is to be ascribed more to incidents of travel than to essential difference of origin.

The directivity ratio, as a percentage, has extremes lying between 20 per cent. and 32 per cent. It is again well represented by a non-periodic and a twelve-month term alone.

The non-periodic term is greater at 13 h. than at 16 h., at 16 h. than at 7 h. The annual term, with its maximum somewhat variable, but roughly midway between the winter solstice and the vernal equinox, has an amplitude greatest at 7 h., least at 13 h.; shorter period terms are unimportant at 7 h., increasingly prominent towards 16 h. (see Table IX).

XII. *Solar Relations.*—The simplicity of the form of the annual variations, the magnitude of the twelve-month term in the expansions, and the incidence of the maxima—a solstitial maximum and equinoctial minimum in intensity, solstitial maxima in deviation of direction of arrival—point very clearly to a solar control of the phenomena under consideration. The crudities of existing methods of attack would seem to preclude the drawing of definite conclusions, at present, as to the nature of this control, whether direct or indirect.

Coefficients of correlation between monthly mean direction of arrival and monthly mean solar altitude at the time of observation amount to $0{\cdot}93 \pm 0{\cdot}03$ at 7 h., $0{\cdot}81 \pm 0{\cdot}07$ at 13 h., $0{\cdot}92 \pm 0.03$ at 16 h., while the coefficients for all observations collected into one single group amount to $0{\cdot}69 \pm 0{\cdot}06$. These

Table IX.

Element.	Reference.		Fourier coefficients.							Phase angles.					Approximate date of maximum, nearest to June 21.				
	Table.	Fig.	C_0.	C_1.	C_2.	C_3.	C_4.	C_5.	C_6.	X_1.	X_2.	X_3.	X_4.	X_5.	1.	2.	3.	4.	5.
Most frequent direction of minimum, 7 h.	I	3	82·33	54·0	21·3	10·3	17·2	30·4	+16·5	228	110	3	83	48	Feb. 2	April 15	July 19	April 20	May 21
Mean direction of minimum, 7 h.	II	4	88·67	52·5	24·1	10·2	3·6	20·6	+4·7	256	85	52	86	44	Jan. 4	May 10	May 29	July 17	May 25
Most frequent direction of minimum, 13 h.	I	3	55·0	22·4	8·1	6·9	4·9	4·2	−5·2	250	117	63	75	26	Jan. 10	April 8	May 18	July 29	June 13
Mean direction of minimum, 13 h.	II	4	52·33	31·6	10·6	7·0	11·8	6·0	+0·33	249	141	82	87	21	Jan. 11	Sept. 14	April 28	July 17	June 18
Most frequent direction of minimum, 16 h.	I	3	69·58	36·6	10·7	9·6	13·5	4·3	+0·7	266	91	80	19	51	Dec. 24	May 5	April 30	June 25	May 18
Mean direction of minimum, 16 h.	II	4	69·17	35·9	9·9	8·0	8·7	5·4	−1·0	262	117	70	9	56	Dec. 29	April 8	May 11	July 5	May 13
Most frequent direction of minimum, total	I	5	62·42	34·3	7·1	13·7	7·9	9·1	+1·7	254	69	79	19	51	Jan. 6	May 27	May 1	June 25	May 18
Mean direction of minimum, total	II	5	64·17	35·6	1·9	12·4	7·7	10·0	+3·2	248	61	77	7	52	Jan. 12	June 4	May 3	July 7	May 17
Mean number at instantaneous minimum— 7 h.	III	7	15·77	2·16	0·59	0·66	0·41	0·44	+0·17	333	35	50	24	7	Oct. 16	July 1	June 1	June 20	July 2
13 h.	III	7	16·76	3·84	2·47	1·25	0·39	0·67	−0·23	20	41	35	41	13	Sept. 1	June 25	June 16	June 2	June 26
16 h.	III	7	17·12	3·97	2·23	1·33	0·62	0·55	−0·20	21	47	37	12	5	Aug. 31	June 19	June 14	July 2	July 4
Total	III	8	16·54	3·14	1·77	1·03	0·33	0·51	−0·09	11	43	39	23	9	Sept. 10	June 23	June 12	June 21	June 30

Mean number at instantaneous maximum—																			
7 h.	IV	7	21·07	3·21	1·55	0·64	0·63	0·61	+0·19	308	35	55	23	8	Nov. 12	July 1	May 26	June 21	July 1
13 h.	IV	7	22·93	4·24	3·80	1·36	0·48	0·79	−0·39	17	43	33	43	15	Sept. 3	June 23	June 18	May 31	June 24
16 h.	IV	7	22·98	4·15	3·43	1·36	0·48	0·64	−0·35	20	50	38	17	2	Sept. 1	June 16	June 13	June 27	July 8
Total	IV	8	22·25	3·23	2·80	1·12	0·51	0·66	−0·10	358	46	42	29	12	Sept. 23	June 20	June 9	June 15	June 27
Mean instantaneous directivity ratio—																			
7 h.	VII	6	24·76	4·32	2·70	0·77	0·07	0·46	+0·09	248	33	96	83	12	Jan. 12	July 3	Aug. 14	July 20	June 27
13 h.	VII	6	26·96	3·83	1·40	1·17	0·17	0·89	−0·13	209	67	101	23	34	Feb. 20	May 30	Aug. 9	June 21	June 5
16 h.	VII	6	25·84	3·96	1·35	1·42	1·16	0·55	−0·19	200	65	95	49	52	Mar. 1	May 31	Aug. 15	May 25	May 17
Total	VII	6	25·97	3·77	1·43	1·21	0·25	0·33	−0·20	218	48	100	48	39	Feb. 11	June 18	Aug. 9	May 26	May 30
Mean number at mean minimum—																			
7 h.	V	—	17·58	2·30	1·14	0·20	0·66	0·63	+0·13	325	39	52	27	8	Oct. 24	June 27	May 29	June 17	July 1
13 h.	V	—	18·22	4·06	3·03	0·91	0·90	0·73	−0·18	14	42	38	35	8	Sept. 7	June 24	June 13	June 9	July 1
16 h.	V	—	18·97	4·26	2·96	0·87	0·69	0·46	+0·07	17	45	31	29	3	Sept. 4	June 21	June 20	June 15	July 6
Total	V	8	18·68	3·61	2·37	0·88	0·76	0·46	+0·05	0	47	33	31	4	Sept. 21	June 19	June 18	June 13	July 5
Mean number at mean maximum—																			
7 h.	VI	—	20·55	3·36	0·83	1·28	0·98	0·97	−0·10	317	27	56	21	6	Nov. 1	July 10	May 25	June 23	July 3
13 h.	VI	—	22·36	4·29	3·26	1·42	0·37	0·63	−0·53	20	43	35	50	17	Sept. 1	June 23	June 16	May 24	June 22
16 h.	VI	—	22·22	3·86	3·45	0·70	0·27	0·43	−0·10	29	45	41	41	13	Aug. 22	June 21	June 10	June 3	June 26
Total	VI	8	21·53	3·20	2·56	0·79	0·33	0·92	−0·35	7	43	42	37	15	Sept. 14	June 23	June 9	June 7	June 24

coefficients are not exact, in that no allowance was made for the interdependence of adjacent values of the variables, but their magnitude is impressive.

The curve of monthly provisional sunspot numbers assigned by Wolfer for the months November, 1920, to September, 1921, bears no marked resemblance to any of the figures of atmospheric data. It may, however, be significant that the curve of sunspot numbers obtained from a decade does bear a notable resemblance to that of monthly mean intensity of disturbance. The curves referring to the epoch 1890-1900 form Table LXVIII of Chree's 'Terrestrial Magnetism.' The curve there given for magnetically "quiet" days seems merely to over-emphasise, as compared with the curve for "all days," the features of agreement between the two groups of data, the March minimum, the June maximum, and the maximum about September.

XIII. *Meteorological Relations.*—Since temperature is a lagging function of solar altitude, correlation coefficients have also been extracted between the mean and modal monthly values and the mean surface temperatures at the time of observation. These coefficients are still high, but lower than those for solar altitudes. The coefficient of correlation between mean number and surface temperature is, for all hours, $0 \cdot 42 \pm 0 \cdot 09$.

Some evidence has been found for a connection between number and atmospheric pressure, but this investigation is not yet complete.

XIV. *Terrestrial Magnetic and Atmospheric Electric Relations.*—A preliminary search has been made for relations between the data under discussion and the outstanding features of data of terrestrial magnetism and atmospheric potential gradient for the same period.

No relation has yet been traced between atmospheric disturbance, whether in respect of (*a*) direction or (*b*) intensity, on the one hand, and (1) the monthly mean inequality of magnetic declination for the hours of observation, (2) magnetic "character" of each day, or (3) electrical character, on the other.

It is, however, to be noted that the annual mean direction of arrival of atmospherics at Aldershot, 151°, is not widely different from the direction of the magnetic meridian, 166°. The periodic variation of the direction of arrival, although vastly greater in amplitude, is in the same direction as that of magnetic declination, and this fact must not be lost sight of in considering the possible common relation of both phenomena to electrical phenomena at very high levels in the atmosphere, the direct magnetic effect being propagated as a rapidly diminishing field of force and compounded with the field of sub-surface origin, while the atmospherics retain, as radiation phenomena, the full range of variation of their source.

XV. "*Intensity Minima*," *Grinders and Clicks.*—On most occasions the observer found it impossible to locate the position of least disturbance by ordinary direction-finding methods, *i.e.*, by estimation of signal strength, but in some 200 cases such a location was possible, and was noted as an "intensity minimum," for comparison with the "count minimum."

These observations fall into two classes; in one class, containing 173 intensity minima, the type of atmospheric did not notably differ from the normal click type. In 162 cases of this class the intensity minimum and count minimum coincided within the limits of experimental error. In one case there was a disagreement of 10°, in two of 20°, in five of 90°; the difference in the remaining three cases was indeterminate, owing to the flatness of the intensity minimum. The disagreements were evenly distributed over the three fixed hours.

The second class is of interest because of the fact that it contains the cases in which a "grinder" distribution, with a recognisable intensity minimum, appeared. This was principally a winter morning phenomenon. In twenty-four 7 h. observations a "grinder minimum" was located by intensity. In nineteen cases the agreement with "count minimum" was close, and all these minima lay between 120° and 140°, eighteen being between 130° and 140°. (All were between December and February.) In five cases there was marked disagreement, four of the five being due to the grinder minimum remaining at 130°—135°, while the count minimum varied to 35°, 70°, and 90°. The fifth case fell in April, when a grinder minimum of 10° was found alongside a count minimum of 100°.

It is a noteworthy phenomenon that a well-marked "grinder" distribution gives an invariable 7 h. minimum of 135° independent of the "click" distribution, that such grinder distributions are confined to the winter months, and that over the whole year they are three times as frequent at 7 h. as at 13 h. or 16 h. Taking these two hours together, seventeen cases of grinder minima were observed (eight at 13 h., nine at 16 h.), and in all cases agreed with the count minimum which, in all save three cases, lay between 65° and 95°.

In one of the 16 h. observations, April 15, 1921, a heavy snow squall was associated with a continuous crepitation of grinder type.

Isolated cases were also noted of a type of atmospheric which was labelled the "beachstone" type, as its sound was very like that of a wave breaking on a pebble beach. This type was noted at 7 h. on February 8, 9, 10, and 28, at 13 h. on January 27, and August 17, and at 16 h. on January 27, 1921.

A further notable point as to grinders is that in all cases when a test was made, the ratio of grinder disturbance to click disturbance was much less on a 2,000 than on a 20,000 metre wave-length.

XVI. *Eclipse Observations.*—Special continuous watch was kept between 7 h. and 19 h. on the day of the solar eclipse of April 8, 1921. The observations are embodied in fig. 9. It will be seen that the direction of maximum disturbance swung steadily clockwise (*i.e.*, in opposition to the normal diurnal variation) between 7 h. and 12 h., and returned to an almost exact agreement with the 7 h. direction before 16 h. The "directivity ratio" in four out of five observations was very markedly greater round the period of maximum eclipse; this was due to an increase, lasting about 2 hours, in the number heard at the most disturbed setting, along with a decrease in the number at the least disturbed setting.

The mean "audibility," measured by the shunted telephone method varied somewhat irregularly throughout the day, but round the eclipse period, 8 h. to 9.30 h., it was markedly less than at 7 h. and 10 h.

XVII. *Working Formula.*—If the year in question be accepted as a fair sample, the following expressions may be taken as working formulæ for the characteristics of atmospheric disturbance to be anticipated at a date D days after June 21.

Direction of Maximum Disturbance.

From true North at 7 h. G.M.T.—$\theta_7 = 175 + 53 \sin (\mathrm{D} + 240)$.
" " at 13 h. " —$\theta_{13} = 143 + 27 \sin (\mathrm{D} + 250)$.
" " at 16 h. " —$\theta_{16} = 159 + 36 \sin (\mathrm{D} + 265)$.
Mean direction for day—$\bar{\theta} = 153 + 35 \sin (\mathrm{D} + 250)$.

Relative Number per Unit Time in Frame Receiver.

(*a*) At instantaneous position of least disturbance—

$$\mathrm{A}_7 = 15{\cdot}8 + 2{\cdot}2 \sin (\mathrm{D} + 333) + 0{\cdot}6 \sin 2 (\mathrm{D} + 35).$$
$$\mathrm{A}_{13} = 16{\cdot}8 + 3{\cdot}8 \sin (\mathrm{D} + 20) + 2{\cdot}5 \sin 2 (\mathrm{D} + 41).$$
$$\mathrm{A}_{16} = 17{\cdot}1 + 4{\cdot}0 \sin (\mathrm{D} + 21) + 2{\cdot}2 \sin 2 (\mathrm{D} + 47).$$
$$\bar{\mathrm{A}} = 16{\cdot}5 + 3{\cdot}1 \sin (\mathrm{D} + 11) + 1{\cdot}8 \sin 2 (\mathrm{D} + 43).$$

(*b*) At instantaneous position of greatest disturbance—

$$\mathrm{B}_7 = 21{\cdot}1 + 3{\cdot}2 \sin (\mathrm{D} + 308) + 1{\cdot}6 \sin 2 (\mathrm{D} + 35).$$
$$\mathrm{B}_{13} = 22{\cdot}9 + 4{\cdot}2 \sin (\mathrm{D} + 17) + 3{\cdot}8 \sin 2 (\mathrm{D} + 43).$$
$$\mathrm{B}_{16} = 23{\cdot}0 + 4{\cdot}2 \sin (\mathrm{D} + 20) + 3{\cdot}4 \sin 2 (\mathrm{D} + 50).$$
$$\bar{\mathrm{B}} = 22{\cdot}3 + 3{\cdot}2 \sin (\mathrm{D} + 358) + 2{\cdot}8 \sin 2 (\mathrm{D} + 46).$$

XVIII. *Conclusion.*—The present paper does not purport to go beyond the presentation of consistent data condensed and analysed, for a test year, of a type of geophysical phenomenon which has not previously been brought into satisfactory quantitative form.

The results are presented now in order that they may be at once available to radio-telegraphists, on the one hand, as a definite set of conditions to be faced in the problem of eliminating atmospherics from receiving apparatus, and to geophysicists as a preliminary contribution to that mass of ordered material on which they must base a theory of the origin, nature, and propagation of naturally occurring electro-magnetic waves.

The data discussed are, admittedly, crude and inadequate. Existing direction-finding methods are not equal to the task of providing satisfactory data, and the limitation imposed by aural reception prevented more frequent observations. No satisfactory method of dealing with the intensities of the individual atmospherics, as opposed to their numbers per unit time, presented itself. "Audibilities," by the shunted telephone method, were measured throughout the greater part of the year, but, mainly because of variation in amplification, did not present sufficient features of regularity to justify their analysis; monthly means are given in Table VIII. The "audibility" is here defined as the shunting resistance required to reduce the number of atmospherics heard per unit time to 25 per cent. of the unshunted number.

(It is hoped that the analysis of records from a directional and relatively quantitative recorder of mean atmospheric disturbance, now in operation at Aldershot, may remove many of the defects of the present series of observations.)

The examination of the data is also at present restricted by the unavailability of much of the geophysical data with which comparisons should be made.

Nevertheless it is satisfactory that, from such data, for a period of a single year, smooth curves adequately represented by, at most, three periodic terms in a Fourier series, should have emerged, and these curves form a justification of the use of existing methods, pending progress, to provide preliminary data in the almost wholly unexplored territory of atmospherics research.

The diurnal and seasonal variations in mean direction of maximum disturbance have been shown to be consistent, an increase of solar altitude bringing with it a counter-clockwise swing of the disturbance vector, with a diurnal range (during the part of the day sampled) of the order of 60°, and a seasonal range of about 100°.

The intensity of disturbance, in so far as it is measured by the number of atmospherics heard per unit time in a receiver held as nearly as possible constant in amplification, is also regular in variation, despite the fact that preliminary efforts to correlate it with other geophysical elements have been less successful than in the case of direction of arrival. The diurnal variation is in agreement with elementary generalisation as to relative disturbance in light and dark, and with the increase of disturbance with conditions favouring thunderstorm formation during daylight.

The work dealt with in this paper was begun under the auspices of the Meteorological Office, Air Ministry, and continued under the Radio Research Board, by whose permission it is now published. Thanks are due to the Superintendent, Kew Observatory, for magnetic and atmospheric electric data, supplied in advance of publication, and to Mrs. R. A. Watson Watt, for much help in the reduction of the data.

The Quantum Theory and Electromagnetic Phenomena.

By William Wilson, Ph.D., D.Sc., Professor of Physics in the University of London.

(Communicated by Prof. J. W. Nicholson, F.R.S. Received August 21, 1922.)

The principles of the quantum are, as is well known, not compatible with classical dynamics and electrodynamics, nor, indeed, with the extended dynamics and electrodynamics of the theory of relativity, although they do not appear to violate the principles of the theory of relativity themselves. The chief hypothesis of the quantum theory, and the one which is most at variance with classical theories, is the hypothesis of stationary states, and this is so well established by experimental evidence that it can hardly be doubted that some parts of the classical electrodynamics, at any rate, will have to be given up or modified. In the present communication, a more general form of the postulate, put forward by the writer some years ago, will be described,* which facilitates the application of the quantum theory to problems such as that of the Zeeman phenomenon, involving magnetic fields, and restricts the region within which the classical and quantum theories are in conflict.

The postulates of the quantum theory, in its present state, are as follows:—

(i) Atoms, or other systems contemplated by the quantum theory, can only exist normally in one or other of certain states called stationary states. While in one of its stationary states the atom does not emit or absorb energy. This is the chief postulate of Bohr's theory of spectra.

(ii) Emission or absorption of energy only occurs during transitions from one stationary state to another, and when, in consequence of such a transition, energy is radiated from the atom, the process is subject to the condition

---

* W. Wilson, 'Phil. Mag.,' vol. 29, p. 795 (1915); vol. 31, p. 156 (1916).

$$E'-E=h\nu, \tag{1}$$

where E′ and E represent the energy before and after the transition respectively, ν is the frequency of the emitted radiation, and h is a universal constant (Planck's constant). This is Bohr's second postulate.

(iii) If the positional co-ordinates, q_s, of the atom or other system are suitably chosen, each q passes periodically through the same range of values, and each integral

$$I_s=\int p_s\,dq_s,$$

where the integration is extended over the complete period of q_s, is a constant, while the atom remains in the same stationary state. The third postulate subjects the stationary states to the condition

$$I_s=\int p_s\,dq_s=n_sh,^* \tag{2}$$
$$s=1,2,3,\ldots f,$$

where the n_s are positive integers and f is the number of degrees of freedom of the system. This postulate has been applied with great success by Sommerfeld in the investigation of the fine structure of the lines of Balmer's series and by Epstein in the investigation of the Stark effect.

The third postulate can be applied without difficulty in cases where (as in the Stark effect) the atom can be treated as a conservative mechanical system of such a kind that the co-ordinates can be chosen in only one way, so that the Hamiltonian principal function S has the form

$$S=S_1+S_2+\ldots+S_f-Et, \tag{3}$$

S_i being a function of q_i and constants, and not containing the remaining q's. We then have

$$p_i=dS_i/dq_i,$$

and each p has the form

$$p_i=\sqrt{(f_i[q_i,\alpha_1,\alpha_2,\ldots\alpha_f])},$$

where the α's are constants, which may conveniently be so chosen that one of them, say α_1, is the energy of the system. Each co-ordinate, q_i, will in general librate between fixed limits, q_{i1} and q_{i2}, which are roots of the equation

$$f_i(q_i,\alpha_1,\alpha_2,\ldots\alpha_f)=0,$$

and the integrals, I, referred to in (2) above are given by

$$I_s=2\int_{q_{s1}}^{q_{s2}}\sqrt{(f_s[q_s,\alpha_1,\alpha_2,\ldots\alpha_f])}\,dq_s. \tag{4}$$

---

* W. Wilson, *loc. cit.*, above.

From equation (4) we get

$$I_i = F_i(\alpha_1, \alpha_2, \ldots \alpha_f),$$

and since there are as many of these equations as the system has degrees of freedom, each of the constants α, and among them therefore the energy of the system, can be expressed as a function of the I's.

Bohr's second postulate, $h\nu = \Delta E$, can be put in the form

$$h\nu = \sum_i \Delta I_i \, \Delta E_i / \Delta I_i,$$

where ΔE_i is the change in the energy of the system corresponding to a transition in which only the quantum function I_i changes, and since

$$\Delta I_i = s_i h,$$

where s_i is an integer, we have for the frequency of the emitted radiation the general expression

$$\nu = \sum_i s_i \, \Delta E_i / \Delta I_i, \tag{5}$$

while the actual frequencies into which the motion of the system can be resolved, and therefore the frequencies of the radiation it would be capable of emitting according to the classical theory, are given by the general formula

$$\nu = \sum_i s_i \, \partial E_i / \partial I_i. \tag{5a}$$

We see, therefore, that to each frequency deduced from the quantum theory, there will correspond, *in general*, a frequency deducible from the classical theory. This fact is the basis of the important correspondence principle of Prof. Bohr. The quantum theory has often predicted frequencies in the emitted radiation which are not actually observed. This is due to the fact that there is nothing in the postulates described above which enables us to determine what values of the integers s_i in (5) are admissible. In this connexion, the correspondence principle has proved extremely valuable. According to this principle, only those frequencies in (5) are admissible to which there is a corresponding frequency given by (5*a*). The principle also furnishes the only means we possess of accounting for the polarisation and intensity of the emitted radiation.

When we study the motion of a charged particle in an electromagnetic field, from the point of view of the general theory of relativity, we are naturally led to regard the quantum statements (2) in postulate iii above as special cases of more general statements which will be given below. The equations of motion of an uncharged particle in a gravitational field are as follows:—

$$\mu \, d^2q^s/dt^2 + \mu \, \Gamma_{mn}{}^s \, dq^m/dt \, dq^n/dt = 0, \tag{6}$$

where Γ_{mn}^{s} is used to represent a well-known Christoffel expression, μ is the mass of the particle (at rest), t is the proper time, and the indices s, m, n are used in accordance with conventions usual in the general theory of relativity, one of these being the omission of the symbol of summation when an index, *e.g.*, m and n above, appears both in a co-variant and in a contra-variant position. Equation (6) may be expressed in the form*

$$dp_s/dt + p_m p_n/2\mu\, \partial g^{mn}/\partial q^s = 0,$$

where

$$p_s = \mu g_{st}\, dq^t/dt,$$

and since

$$p_m p_n g^{mn}/2\mu = \mu g_{mn}/2\, dq^m/dt\, dq^n/dt = \text{constant},$$

we have

$$dp_s/dt - g^{mn}/2\mu\, \partial/\partial p^s\, (p_m p_n) = 0,$$

which reduces to

$$(\partial p_s/\partial q^t - \partial p^t/\partial q^s)\, dq^t/dt = 0. \qquad (7)$$

If the particle has a charge, e, its equations of motion will be given by replacing the right-hand member of (7) by

$$[\partial(eA_t)/\partial q^s - \partial(eA_s)/\partial q^t]\, dq^t/dt,\dagger$$

where A_1, A_2, etc., are the components of the 4-vector potential (generalised to correspond to the co-ordinates q). This expression is simply a generalisation of the well-known expression for the force on a charged particle, namely,

$$e(E + [vH]).$$

We have therefore for the equations of motion of a charged particle

$$(\partial \Pi_s/\partial q^t - \partial \Pi_t/\partial q^s)\, dq^t/dt = 0, \qquad (7a)$$

where

$$\Pi_s = p_s + eA_s.$$

These equations are equivalent to the statement

$$\delta \int_1^2 \Pi_s dq^s = 0.$$

Since in the extended equations (7a) we find p_s replaced by $p_s + eA_s = \Pi_s$, the suggestion naturally arises that p_s in the quantum conditions (2) above should be replaced by Π_s, so that postulate iii now becomes

$$\int \Pi_s dq^s = n_s h, \qquad (2a)$$

$$(s = 1, 2 \dots f)$$

where the appearance of the index s in both a co-variant and a contra-variant position *does not now imply summation.*

---

* W. Wilson, 'Proc. Phys. Soc.,' vol. 31, p. 69 (1919).

† The symbol of summation is omitted in accordance with the convention described above.

The suggestion contained in (2*a*) may be illustrated by applying it to the Zeeman phenomenon in the case of a simple nuclear atom with a massive nucleus. If we refer the system to rectangular axes of co-ordinates with the nucleus at the origin, and suppose a uniform magnetic field, whose intensity is H in ordinary electro-magnetic units, parallel to the z axis, then A_x, A_y, A_z, the x, y, and z components of the vector potential, are

$$A_x = -\tfrac{1}{2}Hy, \qquad A_y = +\tfrac{1}{2}Hx, \qquad A_z = 0,$$

so that the vector A at any point is at right angles to the plane containing the z axis and the point in question. It is convenient to employ polar co-ordinates r, θ and ϕ,

$$x = r\sin\theta\cos\phi, \qquad y = r\sin\theta\sin\phi, \qquad z = r\cos\theta.$$

The components of the vector potential, A, when generalised to suit these co-ordinates, are easily seen to be

$$A_r = 0, \qquad A_\theta = 0, \qquad A_\phi = \tfrac{1}{2}Hr^2\sin^2\theta.$$

Of the equations of motion (7*a*) we only require the third, which, if we neglect the refinements of theory of relativity, is

$$d\Pi_\phi/dt = 0,$$

or

$$p_\phi + eA_\phi = M, \tag{8}$$

where M is a constant. The energy, E, of the system is

$$E = p_r^2/2m + p_\theta^2/2mr^2 + p_\phi^2/2mr^2\sin^2\theta + c^2e\,e'/r, \tag{9}$$

where e and e' are the charges on the electron and on the nucleus respectively in electro-magnetic units, and c is the velocity of light *in vacuo*. If we eliminate p_ϕ from equation (9) by means of (8), and neglect the term $e^2A_\phi^2$ we obtain

$$E = p_r^2/2m + p_\theta^2/2mr^2 + M^2/2mr^2\sin^2\theta - MeH/2m + c^2e\,e'/r,$$

and, therefore,

$$2mr^2E - r^2p_r^2 + MeHr^2 - 2c^2e\,e'mr = p_\theta^2 + M^2/\sin^2\theta,$$

so that

$$\left.\begin{aligned} p_r &= \surd(L/r^2 - 2c^2e\,e'm/r + MeH + 2mE) \\ p_\theta &= \surd(-L - M^2/\sin^2\theta), \\ p_\phi &= M - \tfrac{1}{2}eHr^2\sin^2\theta, \end{aligned}\right\} \tag{10}$$

where L is a constant.

Making use of (2*a*) we get

$$\int\Pi_r dr = \int p_r dr = n_1h; \; \int\Pi_\theta d\theta = \int p_\theta\, d\theta = n_2h; \; \int\Pi_\phi\, d\phi = 2\pi M = n_3h. \tag{11}$$

In the first of these equations the integration is extended from r_1 to r_2 and back again, where r_1 and r_2 are the roots of

$$L/r^2 - 2c^2e\,e'm/r + MeH + 2mE = 0.$$

The integration in the second of these equations is extended from θ_1 to θ_2 and back again, where θ_1 and θ_2 are the roots of the equation

$$-\mathrm{L}-\mathrm{M}^2/\sin^2\theta = 0.$$

The integration of equations (11) gives us

$$\left.\begin{aligned} n_1h &= 2\pi\left\{\frac{c^2ec'm}{\sqrt{(-2m\mathrm{E}-\mathrm{M}e\mathrm{H})}}-\sqrt{(-\mathrm{L})}\right\} \\ n_2h &= 2\pi\{\sqrt{(-\mathrm{L})}-\mathrm{M}\} \\ n_3h &= 2\pi\mathrm{M} \end{aligned}\right\}. \qquad (11a)$$

Eliminating the constants L and M from equations (11*a*) we obtain

$$\mathrm{E} = -\frac{2\Pi^2mc^4e^2e'^2}{(n_1+n_2+n_3)^2h^2}-n_3h\frac{e\mathrm{H}}{4\pi m}. \qquad (12)$$

Now, applying Bohr's second postulate, we have

$$\nu = \frac{2\Pi^2mc^4e^2e'^2}{h^3}\left\{\frac{1}{n^2}-\frac{1}{n'^2}\right\}+(n_3-n_3')\frac{e\mathrm{H}}{4\pi m}. \qquad (13)$$

This gives for the Zeeman separation

$$\Delta\nu = (n_3-n_3')\,e\mathrm{H}/4\pi m.$$

According to the correspondence principle n_3-n_3' is restricted to the values

$$n_3-n_3' = \pm 1 \text{ or } 0,$$

in which case (13) yields the normal Zeeman effect.

This is the result obtained by Sommerfeld* and Debye† by methods similar to that given above, but with the important difference that they employ the quantum restrictions (2) in connexion with co-ordinates which are rotating with the angular velocity $e\mathrm{H}/2m$ about the z axis relatively to the co-ordinates used above.

It will be seen that the stationary states conforming to (2*a*) above, are not in conflict with the electro-magnetic field equations involving curl E and the equation div. H = 0. These do, in fact, become identities when H and E are expressed in the usual way in terms of the vector potential A.

---

* A. Sommerfeld, 'Phys. Zeitschr.,' p. 491 (1916).

† P. Debye, 'Phys. Zeitschr.,' p. 507 (1916).

The Carbon Arc Spectrum in the Extreme Ultra-Violet.

By F. SIMEON, B.Sc., F.Inst.P., Physicist in the Research Laboratory of Adam Hilger, Ltd.

(Communicated by Prof. A. W. Porter, F.R.S. Received September 28, 1922.)

[PLATE 4.]

1. *Introduction.*

Considerable attention has been given in the last few years to the carbon spectrum in the extreme ultra-violet, that from an arc being investigated with a fluorite spectrograph by Wolff* and McLennan, Ainslie and Fuller,* and with a grating spectrograph by McLennan and Lang* and McLennan and Petrie,* while Millikan* has measured the hot-spark spectrum to 360 A.U. On the whole, the agreement of these various workers as to the wave-lengths of the lines is not very good, although McLennan and Petrie are in substantial agreement with Millikan concerning the most intense lines. It was thought worth while therefore to take an opportunity which offered to make another independent measurement of the wave-lengths of a number of lines in the arc spectrum, with the result that some have been photographed in the Lyman region which have been mentioned by no previous worker with the arc, and by the study of some of the lines in higher orders of spectrum, they have been shown in many cases to consist of close groups of lines.

2. *Description of Apparatus.*

The vacuum grating spectrograph used is shown in outline in fig. 1. It consisted of a metal tube with chamber at the side for the introduction of a drying agent. The grating was mounted in a carriage provided with means for giving the grating a limited rotation about each of three mutually perpendicular axes. The carriage could be moved along slides, fixed near to one end of the tube, by a screw, so as to allow of accurately focussing the spectrum upon the photographic plate. This end of the spectrograph was closed by a metal plate. To the other end of the tube was soldered a casting carrying photographic plate and two slits. The position of the plate was fixed, the plate-holder being cylindrical and having the surface against which the plate was held worked to the same radius of curvature as that of the grating. The slits were placed immediately beneath the dark slide and recessed so that each lay on the same focal surface as the grating and photographic plate. Their distance apart was such that a range of spectrum of about 800 A.U. was obtained from either slit, these ranges overlapping by rather more than 80 A.U. for the same setting of the grating, and the focus of the spectrum being equally sharp whichever slit was employed. A shutter, operated by a permanent magnet mounted externally in a rotating collar, enabled two exposures to be made on the photographic plate which measured 6 cm. × 1 cm.

---

* See Table I.

The grating used was a 4-inch Rowland grating of 101·5 cm. radius, with a ruled area of 8·0 cm. × 5·3 cm. and 596 lines per millimetre. As judged by its behaviour in the visible spectrum, it possessed one first order spectrum of considerable brilliancy, but, as will be seen, it was brighter for the visible spectrum than for the extreme ultra-violet.

A water-cooled vacuum arc was used which consisted of a brass tube of about 10 cm. diameter and 20 cm. length, surrounded by another tube of similar length, to form the water-jacket. One end was closed by a brass plate which formed the base, and through this was introduced the positive electrode. A short tube through the circular wall was provided with a flange which fitted the recess of the slit of the spectrograph, while the negative electrode was introduced also through the circular wall and opposite to this tube. Both electrodes were insulated with glass tubing and flexible air-tight joints made with rubber tubing. The electrodes were made of steel rod to which pieces of carbon, or

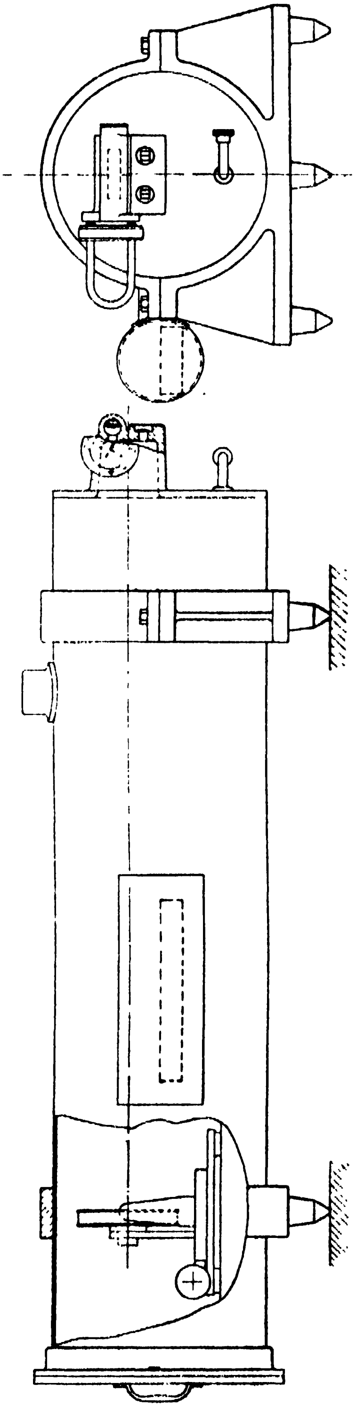


FIG. 1

other material, could be clipped. Another tube nearly opposite to the slit served for evacuation of the arc lamp without this taking place through the slit itself. The glass cover of the arc lamp and the various cover-plates of the spectrograph were made vacuum-tight with laboratory wax.

Trimount vacuum oil pumps alone were used for exhaustion of spectrograph and the lamp. The Schumann plates were generally used "colour-sensitised," as this was found to improve the speed of the plate for the extreme ultra-violet as well as for the ordinary ultra-violet region. Glass plates were used throughout, it being possible to bend thin glass of this width (1 cm.) to the curvature used without breakage.

3. *Experimental.*

The grating was first adjusted so that the 1854 line of the aluminium spark was registered at about 12 mm. from the right-hand end of the plate when the right-hand slit was illuminated. The vacuum arc lamp was then put in position and its exhaust tube connected by a lead pipe to the main body of the spectrograph, so that there was the same pressure at each side of the slit, and hence no tendency for particles disintegrated from the electrodes to be drawn into it. In order to utilise the whole aperture of the grating the positive electrode was mounted eccentrically, so as to be not more than 4 cm. from the slit. It was then found that particles sometimes lodged in the slit, being thrown off from the arc; but not so much trouble was found from carbon as from certain metals such as aluminium. For most of the work a current of about 7½ ampères, taken from the 220-volt supply mains, was used. The arc when struck was brilliant and almost continuous, but it was found preferable not to separate the electrodes more than about ½ cm., as with longer arcs there was a tendency for the formation of a blue glow throughout the lamp, which in extreme cases seemed to spread to the interior of the spectrograph and to cause fogging of the plate. This glow also seemed particularly favourable to obtaining the carbon-monoxide bands of considerable intensity. To reduce the intensity of these bands as much as possible, the procedure was adopted of giving a series of exposures each not exceeding ½ minute duration, and pumping the spectrograph for 5–10 minutes after each one. While the aggregate exposure in most cases did not exceed 5 minutes, the time taken to obtain a photograph, including complete exhaustion of the spectrograph after introduction of the plate, was of the order of 3 hours. The plate glass cover to the arc lamp was found most convenient for observation and control of the arc during exposure, and also gave ready means of access to the interior for change of electrodes.

A discharge tube was used as vacuum gauge. It consisted of a glass bulb of about 10 cm. diameter containing two electrodes about 5 cm. apart. The discharge was taken from a ¼ K.W. transformer capable of giving 15,000 volts, but used with a resistance in the primary circuit consisting of three large lamps in parallel. When a plentiful supply of phosphorus pentoxide was used in the spectrograph the vacuum attained was such that no visible discharge passed through the tube. The tube was again rendered conducting upon striking the arc, and time was allowed for this new discharge to disappear. The time required diminished with the life of the carbon electrodes.

The procedure indicated sufficed to photograph the spectrum as far as 651 A.U. No attempt was made to wash out the spectrograph with hydrogen or helium, nor was liquid air used. Nevertheless it will be seen by reference to Table I that only three lines of intensity five or greater in the spectrum found by Millikan are missing as far as this limit.

A slit width of approximately 0·1 mm. sufficed to show the 1335 doublet, while examination of the photograph under a microscope suggested that other strong lines were double or triple. The slit was then closed to about 0·03 mm. and the resulting photograph showed clearly the 1657 line as a triplet and the 1561 as a doublet. It was then resolved to examine these lines in the second order spectrum. Accordingly the grating was readjusted so that a range of spectrum including 1931 and 2670 (the second order of 1335) could be photographed, and, by use of the shutter, the iron arc in air was photographed upon the same plate. The whole spectrograph was filled with air at atmospheric pressure while the iron comparison spectrum was photographed. A further readjustment of grating sufficed to photograph 1561 and 1657 in the second order, again with the iron comparison spectrum. This photograph revealed the fact that the 1657 line is really a pair of doublets, the centre lines being exceedingly close. As confirmation of this it was decided to proceed to the third order of this line, and the grating was once more adjusted to critical focus for lines in the green-blue of the spectrum. In consequence of the departure of the plate from the focal circle in this position of the grating it was not attempted to photograph more than this line in this order. The departure had not been discernible in photographing the second order.

A reproduction, twice actual size, of a negative obtained is given in fig. 2 (Plate 4). This shows the carbon arc spectrum between 2000 A.U. and 1170 A.U., taken with a slit width of 0·1 mm.

The carbon used for the electrodes consisted of small pieces cut from rods commercially obtainable for arc lamps. To obtain some idea of their purity, the spectrum of the arc in air between two pieces which had previously been used in the vacuum arc lamp was photographed, using a large quartz spectrograph (the Hilger E. 3), and compared with that given by two carbon rods

Table I.—Carbon Spectrum from 250 $\mu\mu$.

1. Eder and Valenta, 'Atlas Typischer Spektren,' 1911. Spark in Air. Wave-lengths corrected to *vacuo* by author.
2. Wolff, 'Ann. der Physik,' vols. 4, 42, pp. 825–839 (1913). Vacuum Arc. Fluorite spectrograph.
3. McLennan, Ainslie and Fuller, 'Roy. Soc. Proc.,' A, vol. 95, pp. 316–332 (1919). Vacuum Arc. Fluorite spectrograph.
4. McLennan and Lang, 'Roy. Soc. Proc.,' A, vol. 95, pp. 258–273 (1919). Vacuum Arc. Grating spectrograph.
5. McLennan and Petrie, 'Trans. R.S.C.,' vol. 15, pp. 15–29 (1921). Vacuum Arc. Grating spectrograph.
6. Millikan, 'Astrop. Journ.,' vol. 52, pp. 47–64 (July, 1920); 'Astrop. Journ.,' vol. 53, pp. 150–160 (March, 1921). Hot Spark. Grating spectrograph.
7. Author. Vacuum Arc. Grating spectrograph.

1.		2.		3.		4.		5.		6.		7.		
λ.	*i.*	λ.	*i.*	λ.	*i.*	λ.	*i.*	λ.	*i.*	λ.	*i.*	λ.	*i.*	$\nu = 1/\lambda \times 10^8$.
2512·83	3											2512·44	3	39802
2509·91	2											2509·47	2	39849
2479·39	15											2479·05	4	40338
				2307·5	2									
2297·65	4			2298·0	9							2297·14	2	43532
				2219·0	2									
				2088·5	5									
						2023	5							
						2020	1							
						1955	1	1955·0	2					
1931·27	2													
1930·76	8	1931·05	8	1930·5	15	1933	25	1931·6	12	1931·1	7	1930·61	5	51797
										1910·2	0			
		1826·30	1							1827·3	1			
		1820·47	2											
		1807·39	3			1813	1	1808·5	3			1807·7*	1	55319
				1758·1	9	1767	1					1760·6	2	56799
				1749·7	5	1759	1			1752·3	2	1751·8	2	57084
		1656·44	6	1656·9	10	1667	25	1657·5	14	1657·6	5	1657·86	8	60319
		1654·96	10									1657·20	8	60341
		1653·94	5									1656·81	8	60357
												1656·12	8	60382
										1624·3	1 B ?			
						1578	1			1577·6	1			

			1562·0	9	1574	25	1561·3	15	1561·3	5	1561·32	9	64048
			1561·2	9							1560·67	9	64075
			1560·5	9							1560·16	9	64096
			1550·7	2					1550·9	3	1550·8	2	64483
			1548·5	3					1548·8	4	1548·3	3	64587
			1482·8	5	1497	1			1482·1	1	1481·7	3	67490
											1467·4	2	68148
			1464·5	6	1477	2			1463·7	2	1463·3	4	68339
											1459·1	2	68535
					1447	2			1432·2	1	1431·6	2	69852
									1426·9	1			
					1416	1			1402·9	4 Si, Ca ?			
									1393·9	4 Si, Ca ?			
									1362·6	5 Ca ?			
									1356·2	1	1355·7	1	73763
					1353	30	1335·3	10	1335·0	15	1335·66	10	74869
					(doublet)						1334·44	10	74938
											1329·60	8	75211
					1348	25	1329·6	2			1329·07	8	75241
					1343	2	1324·1	1	1323·7	7	1323·79	7	75541
									1322·3	2			
					1328	2			1310·5	1	1309·1	2	76388
					1299	3			1299·2	3 Si, Ca ?	1298·8	1	76994
					1297	5			1296·8	2			
									1294·9	2 Si ?			
					1281	4					1280·3*	2	78107
							1277·7	1	1278·7	5	1277·32	4	78289
											1274·3	2	78474
											1264·6	1	79076
							1262·3	1	1262·4	5	1261·21	3	79289
											1260·48	3	79335
									1247·5	7	1247·2	3	80180
									1230·2	2			
											1215·53‡	3	82269
											1209·94	2	82651
									1206·6	5 Ca ?	1206·4	1	82891
											1197·2*	1	83528
									1194·1	3	1194·4	1	83724
											1193·2	1	83808
											1190·3*	1	84012

Table I—(*continued*).

1.		2.		3.		4.		5.		6.		7.		
λ.	*i.*	λ.	*i.*	λ.	*i.*	λ.	*i.*	λ.	*i.*	λ.	*i.*	λ.	*i.*	$\nu = 1/\lambda \times 10^8$.
						1200	1	1176·8	6	1175·6	15	1176·08	7	85028
												1175·56	7	85066
												1174·74	7	85125
										1141·5	4			
										1137·4	3			
										1111·3	2			
										1109·6	3 Si, Ca ?			
										1092·6	3			
										1085·3	3 O, N ?			
										1066·0	8 Si ?	1066·3	1	93782
						1064	2	1037·3	1	1036·7	12	1036·84	5	96447
												1036·22	5	96505
										1022·8	1			
						1039	1			1010·2	10	1010·09	4	99001
										991·1	2 O, N ?			
						1004	3	977·9	1	977·1	12	976·83	5	102372
										966·6	0			
										960·6	0			
										954·4	0			
										945·6	4	945·4	1	105775
										936·4	1			
						934	2	904·7	1	904·1	10	903·7	4	110656
						918	1							
										884·8	1			
										858·5	5	858·3	1	116509
										848·4	0			
										810·0	5			
										806·7	6			
										799·9	5			
										786·5	1			
										749·6	0, O ?			
										743·6	0			

										711·0	0			
										687·3	8	687·2	1	145518
										661·5	0			
										651·5	6	{652·1	1	153351
												651·0	1	153610
										641·8	5	(640·1†)	2	156226
										636·3	3			
										609·5	1, O ?			
										600·2	1, O ?	(598·6†)	1	167056
										595·1	5	(595·1†)	1	168039
										585·7	3			
						584	1							
										574·5	6			
										564·7	3			
										560·5	3			
										549·6	2			
										543·5	2			
										538·4	7			
										533·3	2			
										530·3	2			
										517·6	1			
										511·7	1			
										499·7	4			
										493·7	1			
										459·7	6			
										450·9	1			
										419·8	1			
										386·4	4			
										384·4	4			
										372·1	2 Na ?			
										360·5	0			

---

* Doubtful second order line.

† Observed in second order only (*c.f.* *).

‡ Hydrogen. First member of Lyman series.

which had been chemically treated, obtained from the British Scientific Instrument Research Association. The only line distinguishable in the former spectrum which did not occur in the latter had a wave-length of 2537 A.U., approximately. This is believed to be due to an accidental contamination with mercury in the laboratory subsequent to the use of the carbon poles in the vacuum arc lamp. This supposition is confirmed by the following observations: firstly, an attempt to confirm the presence of this line by using a spectrograph (the Hilger E. 1) of three times the dispersion of the E. 3 instrument was unsuccessful, probably owing to volatilisation of the mercury contamination in taking the previous spectrum; and, secondly, no trace was found on the vacuum photographs of the very strong mercury line at 1850 A.U. Although absolute purity cannot be claimed for the chemically treated carbon rods used for comparison, no considerable impurity is suspected from the spectrograms taken.

4. *Measurement of Wave-Lengths.*

The wave-lengths given in the Table are based primarily upon the second order photographs with iron arc comparison. The negatives were measured with a photo-measuring micrometer reading to 0·001 mm. and the iron lines identified with the aid of Eder and Valenta's 'Atlas Typischer Spektren.' The values there given were compared with those determined by the Bureau of Standards* whenever these were available. These were appropriate as the grating was in air when the iron arc was photographed. The calculated carbon wave-lengths are, however, vacuum values. The values thus determined for the stronger lines were then used as datum values in the measurement of the first order plates. By slightly varying the setting of the grating the three strong lines in the neighbourhood of 1000 A.U. could be obtained in the range of spectrum from either slit, and the wave-lengths of the shorter lines determined by extrapolation. The values were also checked by Lyman's displaced spectrum method.†

The accuracy of measurement of the second order lines which were compared with the iron spectrum should be considerable. From calculation of the data of the grating, and by measurement of known lines, it was found that the dispersion of the grating was approximately 0·06 mm. per A.U. Except in the case of over-exposed lines it should, therefore, be possible to set rather more accurately than 0·1 A.U. The values for the iron lines are known to greater accuracy than this, so that these lines should be known in the first order to about 0·05 A.U. Assuming perfect adjustment, we should expect corresponding accuracy for remainder of the lines. Departure from perfect adjustment might be occasioned by (1) error of focussing, (2) inclination of the photographic plate to focal circle, (3) displacement of the slit from focal circle, (4) non-coincidence of plate with curve of plateholder, (5) distortion of instrument on evacuation, and (6) incorrect curvature of photographic plate. Now it was found by trial that error of focus of ½ mm. could be easily detected, so that this cause would introduce an error of 0·05 per cent. For the greater number of lines the distance from a datum value would not exceed 200 A.U., so that the amount of error becomes 0·1 A.U. The second and third errors would arise in manufacture of the instrument, and neither would be expected to exceed 1/10 of the error just discussed. The fourth error would depend upon the difference of thickness of plate allowed for in the plate-holder and the actual thickness of plate. This again should not exceed ½ mm. and is probably less. It is conceivable that some change of shape of the spectrograph might be occasioned by removal of the air. This would affect the second order photographs, in which the iron arc comparison was taken with air filling the spectrograph. No evidence at all, however, was obtained of the presence of such an effect. It will be noticed in connection with the sixth error that the actual curvature used was one-half the correct focal curvature. Calculation shows that the consequent error of focus, when averaged along the plate, should be about one-half that actually discernible. Such a variation of focus was not observed in practice. It is to be expected that the sum of these errors should not exceed 0·3 A.U.

---

* Sci. Papers Nos. 251 and 274.

† 'Phys. Rev.,' vol. 16, p. 257 (1903).

The wave-lengths obtained for the 1657 group may be quoted to show the agreement obtained by measurements in different orders of spectrum. In the first column are values measured in the second order by direct comparison with the iron arc spectrum, while those in the second column were deduced from a knowledge of the grating constant and the value of the wave-length of the 2479 line obtained in the first order by comparison with iron lines. The second order of this line falls very close to the third order of 1657.

Measured.		Deduced.
1657·87 A.U.	....	1657·86 A.U.
1657·18	....	1657·21
1656·81	....	1656·81
1656·14	....	1656·10

5. *Results.*

The wave-lengths of the lines observed, and their estimated intensities, are given in the last column of the Table. In the remaining columns are given the values of all other workers known to the writer. With each reference the source and type of spectrograph is given.

The wave-lengths obtained from plates upon which the iron arc was registered are quoted to the second decimal, as it is believed that these should be correct to 0·1 A.U. It would follow that, for lines obtained in the second order, the error should be of the order of 0·05 A.U.

6. *Discussion of Results.*

It will be seen by reference to the foregoing Table that a number of lines have been photographed as close groups of lines. A word may be said concerning some of these groups.

Particular attention has been paid to the 1657 group, an effort to resolve the apparent triplet seen in the first order spectrum having revealed that it really consists of two close pairs of lines, the separation of the centre two lines being approximately one-half that of each other pair of neighbouring lines. The more refrangible pair is distinctly sharper than the other. In fig. 3 (Plate 4) is shown this line as obtained in the third order, together with the second order of the well known 2479 line. The reproduction is a tenfold enlargement of the negative. It is interesting to note that Wolff obtained this line as a triplet in working with a fluorite spectrograph.

The 1561 line appears in the second order spectrum as a group of three, and was so obtained by McLennan, Ainslie and Fuller with a fluorite instrument. This line was not observed in higher orders.

The 1335 pair was noted as a doublet by McLennan and Lang, and the writer agrees with them in considering it as the most intense line in the vacuum arc spectrum of carbon. It was invariably the first to appear during development whenever present.

The pairs of lines at 1329, 1260, 1194, 1036 and 651 A.U. seem to have been observed now for the first time. That at 1036 has been photographed in the third order, together with the second orders of 1657 and 1561. So far it does not show any further complexity. The 651 pair has been observed in the second order only, for the 651 line of the first order, which marks the limit of the present investigation, was not resolved.

The 1175 group presents some difficulties. It was obtained in the second order as a distinct group of three, while on the occasions on which it was resolved in the first order, two only were discernible.

Three single lines, given in Table I, call for comment. These are 1280·25 A.U., 1197·2 A.U., and 1190·3 A.U. The first and last of these are probably the same as those found by Millikan at 1282·0 and 1191·0, and which he showed were respectively the second orders of lines found at 641·8 and

595·1. It would seem, therefore, that these lines are also present in the arc spectrum, although they were not obtained in the first order. The other line is near to the value 1199·8 given by Lyman* for a helium line, and corresponds closely to the second order of the line found by him at 599·0. The first order values are therefore entered in the Table in brackets. The fact that these second order lines were distinctly obtained, while no trace could be found of the corresponding first order lines, indicates that the ruling of the grating was such as to throw considerably more energy in the direction of the longer wave-lengths. This is borne out by two other observations. (1) The time of exposure for the third order of 1657 did not exceed ½ min., while the average exposure for the first order plates was 5 min. Allowing for the fact that the bright lines on these plates were somewhat over-exposed, to bring out the fainter lines, we may take as correct exposure for 1657 in the first order about one-half that actually given, *i.e.*, the spectrum was about five times as bright at 5000 A.U. as it was about 1600 A.U. (2) On removing the grating from the spectrograph and examining it by eye, it was seen that while the first order of the visible spectrum was brilliant, the central image was of a dull reddish hue. The reflecting power for very short wave-lengths in the direction of this central image may be expected to be very small.

In view of the fact that no helium was introduced into the spectrograph, the occurrence of the 599 line in the present work may be taken as an indication that this line should be included among those mentioned by Millikan (*loc. cit.*) and McLennan and Petrie (*loc. cit.*) as attributable to carbon in the helium spectrum found by Lyman (*loc. cit.*).

The line found at 1808 A.U. is retained in the present list of wave-lengths, as it appeared to be composite when photographed with a fine slit. It is thought that a first order line is coincident with the second order of the strong line at 904. This is confirmed by Wolff's value obtained with a fluorite spectrograph.

7. *Summary.*

The results of this investigation may be stated as follow:—

(1) The arc spectrum of carbon has been shown to give lines in the Lyman region at 1194, 945, 858, 687, 651, 640, 599, and 595, which have not been previously observed, and which correspond with prominent lines in the "hot spark" spectrum studied by Millikan.

(2) Groups of lines have been found at 1657, 1560, 1335, 1329, 1260, 1194, 1175, 1036, and 651, of which those at 1329, 1260, 1194, 1036, and 651 do not seem to have been resolved at all by any other worker, and that at 1657 has not been completely resolved heretofore.

---

* Lyman, 'Astrop. Journ.,' vol. 43, pp. 89–102 (1916).

(3) A new determination of wave-length has been made of the prominent lines in the carbon arc spectrum.

In conclusion, I would express my indebtedness to the firm of Adam Hilger, Ltd., for placing at my disposal all the facilities for this investigation and for permission to publish the results. I would also acknowledge the help of my assistant, Mr. Barnes, in the routine experimental work.

Changes in the Charge of an α-Particle passing through Matter.

By G. H. HENDERSON, Ph.D., 1851 Exhibition Scholar of Dalhousie University, Halifax, N.S.

(Communicated by Prof. Sir E. Rutherford, F.R.S. Received October 28, 1922.)

[PLATE 4.]

§ 1. *Introduction.*

It is remarkable that investigations on the velocity of α-particles have hitherto failed to detect with certainty particles of velocity less than about 8×10^8 cm. per second ($0{\cdot}4V_0$, where V_0 is the initial velocity of α-particles from RaC). Geiger* believed he had found α-particles with velocities as low as $0{\cdot}3V_0$. However, later work of Marsden and Taylor† showed that when the velocity of an α-ray beam had fallen to about $0{\cdot}45V_0$, any further increase of absorption caused a rapid decrease in the number of particles; the number fell to zero when the velocity had decreased to about $0{\cdot}4V_0$.

Thus the minimum observed energy was that which would be acquired by the α-particle in falling through about 600,000 volts. It is surprising that particles with such enormous energy should vanish, leaving no trace. Positive rays of much less energy have of course been detected in several ways.

A further difficulty arises in connection with the straggling of a beam of α-particles. The velocity of the beam is at first homogeneous, but towards the end of the range it becomes inhomogeneous and the remaining ranges of the particles are not equal; the beam is said to be straggled out. It has been shown by the writer‡ that the straggling measured experimentally is much greater in extent than that calculated from recognised theory.

* Geiger, 'Roy. Soc. Proc.,' A, vol. 83, p. 505 (1910).

† Marsden and Taylor, 'Roy. Soc. Proc.,' A, vol. 88, p. 443 (1913).

‡ Henderson, 'Phil. Mag.,' vol. 44, p. 42 (1922).

An attempt made by the writer to trace out the cause of these anomalies has led to the discovery of several new phenomena, which will be described in this paper.

§ 2. *Apparatus.*

The results of Geiger, and Marsden and Taylor were all obtained by means of the scintillations produced on a zinc sulphide screen. The scintillations produced by low velocity α-particles are very faint, practically on the threshold of vision. It might be possible that if scintillations were produced by particles of still lower velocities, they would fail to be detected by the eye. On the other hand, the particles might not be able to penetrate sufficiently far into the zinc sulphide to stimulate an appreciable proportion of scintillations.

For these reasons it seemed advisable that some other means of detecting the α-particles be adopted, and as such the photographic method recommended itself. Since these slow particles had small penetrating power it was decided to use Schumann plates. These plates, which were supplied by Adam Hilger, Ltd., were used throughout this investigation, and proved for this purpose more sensitive even than the plates used in positive ray work.

The usual arrangement of line-source, slit and photographic plate in a magnetic field was used for determining the velocities of the α-particles. The strength of field usually employed was 4100 Gauss. The pole pieces of the electro-magnet extended slightly beyond the region between source and plate, ensuring a fairly uniform field. The distances, source-slit and slit-plate were each about 5 cm., but varied slightly in different experiments. These distances were measured to 0·1 mm.

The source, slit and plate were mounted on a brass base which slid into an air-tight brass box, fitting between the pole pieces. The ends of the box were closed by ground glass plates. The vacuum was quickly established and maintained by a Gaede mercury pump to 0·01 mm. or less of mercury.

The source first used consisted of an α-ray tube containing two or three millicuries of radium emanation. The walls of the tube were equivalent in stopping power to between 6 and 7 cm. of air. It was thought that scattering might become a very important factor at low velocities, and such a source would permit particles deflected through large angles to pass through the slit to the plate. As it was found later that practically the same results could be obtained with an active wire and plane absorbing screen, the latter was used for convenience.

Platinum wires were exposed to radium emanation, the source obtained being equivalent in γ-ray activity to between 10 and 20 mgrms. of radium. The diameter of the wire usually used was 0·3 mm., and the width of the slit was also 0·3 mm., though occasionally thicker wires and wider slits were used. The edges of the slit were of thin aluminium to reduce scattering. The absorbing screens were mica, unless otherwise noted, mica being chosen on account of its uniformity. Usually two sheets of gold, each equivalent to about 0·45 mm. of air, were placed between the source and the mica to avoid possible phosphorescence from the source. The stopping power of the mica was determined by weighing, assuming 1·42 mgrms. per square centimetre, equivalent to 1 cm. of air at 15° C. Fine adjustment of the thickness of absorbing material could be secured by tilting the mica sheet about the source as axis.

Immediately in front of the lower half of the photographic plate was placed a mica sheet of 1·8 cm. air equivalent. The purpose of this strip was to show the difference between the effect of the α-rays and that of the γ- and stray β-rays, the former being absorbed by the strip and the latter two scarcely absorbed at all. The effect of this strip may be clearly seen in the lower parts of the photographs reproduced.

§ 3. *The Photographs.*

When photographs were taken with the tube source, a rather diffuse band was seen on the plate, corresponding to the ordinary deflected beam of α-rays. In addition, however, there appeared a faint but clearly distinguishable band about halfway between the usual deflected band and the position of the band in the absence of a magnetic field. The same phenomenon was subsequently obtained with wire sources.

Thereupon photographs were taken systematically under different conditions of absorption of the α-particles from a wire source. The usual procedure was to leave the magnetic field off for one or two minutes after evacuation, so as to yield a fiducial line of the undeflected beam. Evacuation did not begin until RaA had practically disappeared. Thereafter the field was established and kept constant for the remainder of the exposure, about two hours.

It will be well to follow the results in the order of increasing absorption of the beam of α-rays. When no absorbing material was placed over the source (range = 6·95 cm.) the normal deflected and undeflected bands were obtained. No appreciable sign of a band between the two could be noticed. Incidentally opportunity was taken to check the results by determining the velocity of the beam from the constants of the apparatus. For thicknesses of absorbing material ranging from 0 to 6 cm. air equivalent, the velocities agreed within the limits of error (about 2 per cent.) with the results of Marsden and Taylor. For absorption greater than about 6 cm. the band became too diffuse to justify carrying the process further.

We return to the appearance of the photographs. When a sheet of mica of

1·42 cm. air equivalent was placed over the source, emergent range = 5·53 cm., a very faint band appeared between the deflected and undeflected bands. For brevity this band will in future be referred to as the middle band. As more and more absorbing material was placed over the source, further photographs showed the middle band gradually increasing in intensity compared with the normal deflected band.

Some estimate of the intensity of the middle band could be made with the help of the undeflected band. By adjusting the time during which the field was off, the intensities of the middle and undeflected bands could be made roughly the same, while from the decay constants the ratio of the numbers of α-particles in the deflected and undeflected bands was at once known. In the case of 5·53 cm. emergent range, the intensity of the middle band was of the same order of magnitude as that which would be produced by 1 per cent. of the α-particles in the ordinary deflected band. At 4·08 cm. emergent range the ratio of intensities was scarcely increased, while at 0·75 cm. it was of the same order of magnitude as that produced by 7 per cent. of the particles in the deflected band. Great importance should not be attached to these figures, since photographic intensity is often misleading as a measure of the number of particles affecting the plate.

From being a faint and barely distinguishable band, the middle band gradually increased in clearness till at about 1·5 cm. emergent range its appearance was well defined. It then resembled closely the normal deflected band. Both edges were well defined, the inner edge rather sharper than the outer, as in the case of the deflected band. In width it appeared slightly narrower than the deflected band.

An important point to notice is this. When the emergent range was greater than about 1·8 cm., the normal deflected beam was able to penetrate through the mica strip half covering the photographic plate, and this band then extended the full width of the plate. So also did the middle band. When the absorption became so great that the deflected band stopped at the edge of the strip, the middle band did the same.

We turn to the position of the middle band. This band did not seem to be exactly midway between the deflected and undeflected bands when it first made its appearance, but rather nearer the former. This cannot be regarded as established, however, for contrast effects near the strong deflected band were a disturbing factor with the rather small separation of the bands at this velocity. Before this point can be settled, photographs must be taken with an apparatus giving considerably greater separation, so as to render contrast ineffective.

When the emergent range was less than about 2·5 cm. the separation became so great that contrast ceased to be a disturbing factor. The middle band was then exactly midway between deflected and undeflected bands, within the limits of experimental error.

These results may be seen from Table I, which gives measurements on a few typical plates. Each reading (in millimetres) is the mean of three or more independent comparator settings on the lines. The accuracy depended greatly on the nature of the line. On a sharp edge settings could be repeated to 2 or 3 hundredths of a millimetre. The difficulty of setting on the faint middle band in the early stages of absorption has already been referred to. On the other hand, when the emergent range was less than 1 cm. the bands began to become diffuse owing to straggling, and accuracy was again less. Plate No. 25, Table I, was obtained with 0·3 mm. wire and slit, Nos. 20 and 18 with thicker wires and wider slits.

Table I.

Plate No.	Emergent range.	Deflected band.		Middle band.		Undeflected band.		Mean distance.	
		Outer edge.	Inner edge.	Outer edge.	Inner edge.	Outer edge.	Inner edge.	Deflected to middle band.	Middle to undeflected band.
	cm.	mm.	mm.	mm.	mm.	mm.	mm.	mm.	mm.
25	3·35	4·01	3·34	2·36?	1·92?	0·45	0·00	1·53	1·92
20	1·30	5·88	4·30	3·62	2·23	1·65	0·00	2·16	2·10
18	0·72	6·93	5·33	4·01	2·86	1·15	0·00	2·69	2·86

§ 4. *Photographs with Low Velocity Particles.*

When the emergent range was less than 1 cm. the character of the bands began to change rapidly. The middle band rapidly increased in intensity relative to the deflected band. It also ceased to remain midway between the other two bands and could be definitely seen to be nearer the deflected band. This came about because the deflected band ceased to move much further outward as the emergent range decreased, while the middle band marched steadily outward. Plate No. 18, Table I, shows the beginning of this process. This photograph is also reproduced in the accompanying reproduction (Plate 4, lower figure). Unfortunately, most of the photographs obtained are not sufficiently intense for reproduction and only a few could be used for this purpose.

At about 3 mm. emergent range the intensity of the middle band had become equal to that of the normal deflected band. Plate No. 11 represents this stage. Soon afterwards the two bands began to merge together, the intensity of the normal deflected band falling off rapidly.

Another interesting phenomenon occurred when the emergent range was reduced to about 4·5 mm. A band appeared in the position of the undeflected band when the field was kept on during the whole of the exposure. On further absorption this band increased relatively in intensity compared to the middle band. It never became as strong as the middle band, however, and seemed more readily absorbed.

The bands disappeared altogether from the plate, when the thickness of mica was increased to about 11 mgrm. per square centimetre. If the stopping power of the mica remained constant, this thickness would be equivalent to 7·7 cm. of air. As the stopping power decreases near the end of the range, this mica would probably be equivalent to about 7·2 cm. of air. For this reason the air equivalent of foils thicker than 6·6 cm. is a little uncertain. Plate 14 shows the state of affairs where the emergent range was probably of the order of 1 mm. It was impossible to secure satisfactory photographs with smaller emergent range.

We have now traced out in some detail the "life-history" of a beam of α-particles, as revealed by the photographs. Before going on to a discussion of the results, some further experiments will be described.

A wedge was made up of sheets of gold-leaf, increasing in thickness from one to ten sheets of gold-leaf (0·4 to 4 mm. air equivalent). This wedge, together with an additional aluminium foil of 1 mm. equivalent, was placed directly over the photographic plate, replacing the mica strip previously referred to, with its edge parallel to the length of the plate. A photograph was then taken, when the emergent range from the source was 3·3 mm. In the absence of the wedge, the intensities of the middle and the deflected bands would have been about equal. With the wedge, both bands faded away towards the thicker side of the wedge. The middle band appeared to fall off a little more rapidly than the normal deflected band. This photograph (No. 31) has also been reproduced in the Plate.

To see if different results might be obtained with absorbing material other than mica, photographs were taken when gold was the absorbing material. Composite gold foils, made up of 104 and 128 single leaves, were used, the emergent ranges being 2·40 and 1·35 cm. respectively. So far as could be judged from a photograph, the same results were obtained as with mica. However, this comparison is only a rough one.

A photograph was taken while air at a pressure of about 2 mm. was retained in the chamber. The emergent range was 1·35 cm. The middle band had completely disappeared, although with a good vacuum this band would have been quite distinct. The air in the path of the beam would have been equivalent to about 0·3 mm. at normal pressure, and hence would have caused little change in the appearance of the plate if concentrated at the surface of the source.

Finally, a photograph was taken when the slit was covered with two thicknesses of gold-leaf (total = 0·9 mm. air equivalent), and the emergent range was 1·35 cm. The usual vacuum was maintained. The middle band was again found to be absent, while the deflected band appeared more diffuse.

§ 5. *Discussion.*

These photographs disclose a number of new features which require explanation. The so-called middle band has been followed through a variety of experimental conditions. From its general behaviour, it seems fairly clear that this band must be due to positively charged particles of atomic mass, moving with great velocity. They are deflected by a magnetic field, with even more difficulty than α-particles. They appear to have a more or less definite range, since they are able to pass through a strip of mica when their velocity is high and are stopped by it below a certain velocity.

The most obvious hypothesis that may be made is that this band is due to singly-charged α-particles. Though it is not claimed that this hypothesis is proved conclusively by these experiments, all the evidence so far accumulated is consistent with and points towards such a conclusion. At some time or other the α-particle must become degraded into an ordinary helium atom. The question is, at what velocity does it begin to attach electrons ?

If we disregard the position of the middle band on its first appearance, the band lies exactly midway between the deflected and undeflected bands. This is what would be called for by particles of the same momentum and half the charge of the α-particles, and is independent of the constants of the apparatus.

At about 9 mm. emergent range the middle band begins to move from its central position, and lies nearer the deflected band. By this time straggling has become appreciable and the normal deflected band appears diffuse, particularly on the outer edge. It would appear that this edge, due to particles of lower velocity, fades away more rapidly than the inner edge as the emergent range decreases, while the outer edge of the middle band is strengthened at its expense. Thus the centre of the middle band appears to be displaced towards that of the deflected band.

The middle band cannot be produced by recoil atoms from the radioactive source, for these recoil atoms are absorbed by 1/10 mm. of air. Incidentally, no trace of a central band, which might be caused by recoil atoms, was obtained when no absorbing material was over the source. This is in agreement with Wood and Steven,* who failed to find any photographic effect of the recoil atoms on Schumann plates.

The middle band might conceivably be due to atoms from the absorbing screen which, by collision with α-particles, were endowed with the whole momentum of the latter and sent off with a single positive charge. Such an explanation seems very improbable, however, particularly when the homogeneity of the middle band and the considerable penetrating power of the particles causing it are considered.

It is desirable that further experiments should be carried out. Electrostatic deflection of a beam of slow α-rays and a sytematic examination of the region between deflected and undeflected bands by means of scintillations should settle any possible doubt as to the origin of this middle band. Owing to the writer's departure from Cambridge, only the experiments described could be completed. The investigation will be continued.

It was known that scintillations were to be seen over the whole region between the deflected and undeflected bands. These were usually regarded as being due to α-particles scattered by the edges of the slit. Some might also be due to the slight traces of air present in the vessel (see end of § 4). It is thus very difficult to recognise a band superimposed upon such a background without a long and rather tedious examination of the whole region. It is doubtless owing to this difficulty that the middle band has hitherto escaped observation. A photograph, however, readily shows up such a band.

Sir Ernest Rutherford has very kindly informed me that a recent systematic counting of this region showed clear evidence of a maximum at the midway point between deflected and undeflected bands, when the remaining range was 9 mm. This result practically seems to settle any doubt that the middle band is due to singly charged α-particles.

If this view is correct, we may readily explain the appearance of the undeflected band when the magnetic field is on continuously (small emergent range). This is probably due to α-particles which have acquired two electrons and are uncharged; in other words, have become ordinary helium atoms.

Following the hypothesis further, the point at which the α-particle would first pick up a charge would depend on the probability that a certain type of collision would occur with electrons of the absorbing material. There would probably be an interplay of charges, more rapid at lower velocities, and the particle might become, several times in its career, a doubly-charged, a singly-charged, and even an uncharged particle. In a beam of α-particles, at any point along the range, there would be definite proportions of these three types, the proportions being governed by probability considerations.

---

* Wood and Steven, 'Proc. Phys. Soc. London,' vol. 27, p. 189 (1915).

This may account for the disappearance of the middle band when a little air was admitted into the vessel. The small proportion of α-particles starting out from the surface of the absorbing material as singly-charged particles may lose their charge at different points in the body of the vessel, and the middle band may be spread out into the general background of the plate. A similar explanation accounts for the disappearance of the middle band when gold-leaf is placed over the slit, although scattering may become appreciable here.

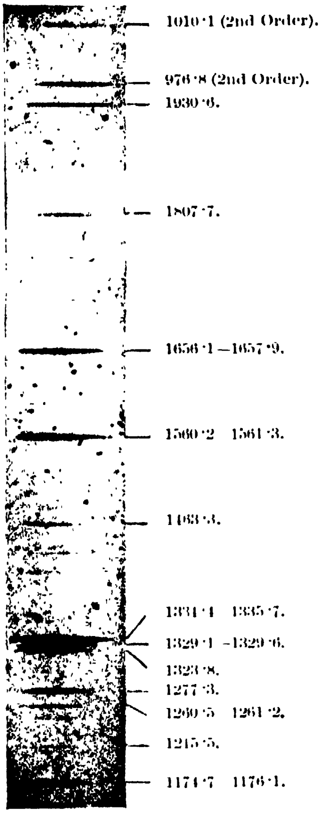
Little can be said as to the penetrating power of the singly-charged particles. The photographs obtained with the gold-foil wedge showed that the middle band was absorbed somewhat more easily than the deflected band. On its passage through the wedge a singly-charged particle probably loses and gains its electron many times. The same idea would apply to a doubly-charged particle; hence the average charged condition of the particles giving rise to the two bands would be different, but not widely different, while passing through the wedge. If this is so, the penetrating power of a singly-charged particle should be very much less than that of a doubly-charged one. Such a conclusion can be only regarded as provisional however.

From the above considerations it appears that the α-particle picks up electrons and becomes singly-charged with some readiness when its energy has fallen to about 800,000 volts, it may do so more rarely at higher energies however. Traubenberg and Hahn,* working on the penetrating power of hydrogen positive rays, found that these particles lost their charge readily at energies less than 7500 volts. On almost any theory this critical energy should be considerably higher for singly-charged helium than for hydrogen, on account of the greater mass and higher ionisation potential of the former. It should be clearly pointed out, however, that little importance should be attached to a comparison between results obtained at widely differing energies. The mechanism involved at 800,000 volts is probably very unlike that involved at energies less than 1/100 of this.

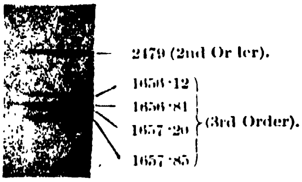
The original object of the experiments was to search for α-particles of velocities less than $0{\cdot}4V_0$. The following may be said of the deflected beam. With high absorption the outer edge of this band became very diffuse, the intensity fading away gradually with no sign of an abrupt ending. It is, of course, impossible under these circumstances to say whether a minimum velocity exists. However, an appreciable darkening of the plate could be observed, at least as far as a point corresponding to a velocity of $0{\cdot}25V_0$.

---

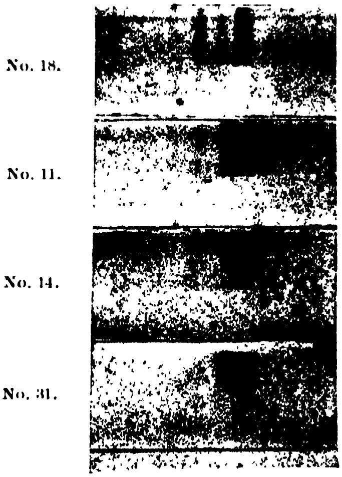
* Traubenberg and Hahn, 'Zeit. für Phys.,' vol. 9, p. 356 (1922).



Simeon Fig. 2.



Simeon—Fig. 3.



Henderson.

If the hypothesis of singly-charged particles be adopted, the minimum velocity appreciable would at least be as low as $0{\cdot}15V_0$. For still lower velocities practically all the particles may be degraded into neutral particles, and further decrease of velocity would be impossible to follow by the method of magnetic deflection.

Finally, a word may be said as to the bearing of these results on the straggling of α-particles. It was pointed out that the straggling near the end of the range of an α-ray beam was much greater than was called for by known probability effects. Evidence is offered in this paper to show that the α-particle readily picks up and loses charges near the end of its range, and that the penetrating power of a singly-charged particle is considerably less than that of a doubly-charged one. As Flamm and Schumann* conjectured some time ago, this process of gaining and losing electrons may introduce a probability variation in the ranges of the individual α-particles sufficient to account completely for this excess straggling.

The results given in this paper have been obtained with α-particles from RaC. It has been shown elsewhere that α-particles from different radioactive sources yield ionisation and scintillation curves which are practically identical near the end of the range. From this it may be inferred that all the results discussed above hold true for any type of α-particles; when their emergent ranges are reduced to the same value by suitable absorption the same results should be found.

Summary.

A beam of α-rays, reduced to a low velocity by absorbing material, was deflected by a magnetic field and studied photographically with Schumann plates. The photographs showed a new band which appeared halfway between an undeflected and a normal deflected beam.

The hypothesis is put forward that this band is due to α-particles which have captured an electron and are singly-charged. The photographs also show a band which is ascribed to neutral α-particles. Further evidence bearing on this hypothesis is presented, and the process by which electrons are captured and lost is discussed. It is pointed out that this hypothesis offers a reasonable explanation of previously outstanding difficulties connected with α-particles near the end of their range.

In conclusion, I wish to express my best thanks to Prof. Sir Ernest Rutherford for his stimulating advice and criticism. I also wish to thank Mr. G. R. Crowe for the preparation of the radio-active sources.

---

* Flamm and Schumann, 'Ann. der Phys.,' vol. 50, p. 655 (1916).

DESCRIPTION OF PHOTOGRAPHS. (Plate 4, Lower Figure.)

(Undeflected beam on left in all photographs.)

No. 18. 7·2 mm. emergent range. Magnetic field off for 10 minutes.
No. 11. 3·3 mm. emergent range. Field on continuously.
No. 14. About 1 mm. emergent range. Field on continuously.
No. 31. 3·3 mm. emergent range. Field on continuously. Mica strip previously over lower half of plate replaced by wedge of gold leaves.

The Crystalline Structure and Properties of Tartaric Acid.

By W. T. Astbury, B.A.

(Communicated by Sir William Bragg, F.R.S. Received October 3, 1922.)

Ever since the time of Biot and Pasteur, tartaric acid has been the centre of most interesting and fruitful discussions among chemists, physicists and crystallographers. It is hoped that this paper will help to throw light on some of the problems with which its name has so long been associated.

1. *Crystalline Structure.*

The X-ray observations of this work were taken with the aid of a Bragg ionisation-spectrometer and a Coolidge bulb of molybdenum anticathode. The crystals examined were of the ordinary (dextro) variety of tartaric acid.

On p. 303, Vol. 3 of P. Groth's "Chemische Krystallographie," the following data are given:—

d- and *l*-tartaric acid, $C_4H_6O_6$.

Specific gravity = 1·759. Monoclinic Sphenoidal.

$$a : b : c = 1{\cdot}2747 : 1 : 1{\cdot}0266\,;\ \beta = 100^\circ\ 17'.$$

First observations with the X-ray spectrometer gave reflections from the (100) face at the following approximate angles of ionisation chamber:—

1st order.	2nd order.
4° 57′	10° 20′

These angles refer to the Kα-line. From them, if θ = glancing angle for first-order reflection, and taking $\lambda = 0{\cdot}714$ A.U.,

$$d_{100} = \lambda/2 \sin 2^\circ\ 42' = 7{\cdot}58 \text{ A.U.}$$

But
$$a = d_{100} \times \operatorname{cosec} 79^\circ 43' = 7{\cdot}70 \text{ A.U.}$$
$$b = 6{\cdot}04 \text{ A.U.}$$
$$c = 6{\cdot}20 \text{ A.U.}$$

∴ Volume of unit monoclinic cell... $= abc \cos 10^\circ\ 17'$ A.U.

Mass of unit monoclinic cell ... $= abc \cos 10^\circ\, 17' \times 1{\cdot}759$ A.U.

Molecular weight of $C_4H_6O_6$ ... $= 150$

$\therefore$ Number of molecules per cell ... $= \dfrac{abc \cos 10^\circ\, 17' \times 1{\cdot}759}{150 \times 1{\cdot}66} = 2{\cdot}006.$

If we take this to mean 2 exactly, we can now calculate the exact values of *a*, *b* and *c* from the crystallographic data. For

$$2 \times 1{\cdot}66 \times 150 = b^3 \times 1{\cdot}2747 \times 1{\cdot}0266 \times \cos 10^\circ\, 17' \times 1{\cdot}759.$$

Whence $a = 7{\cdot}693$ A.U.; $b = 6{\cdot}037$ A.U.; $c = 6{\cdot}195$ A.U.

We are now in a position to calculate the spacings to be expected for other planes on the crystal. The calculation is most simply performed by using the interfacial angles given by Groth, the stereographic projection and the equation to the normal to any plane P whose indices are (h, k, l).

$$\mathrm{OP} = \frac{a}{h}\cos \mathrm{XP} = \frac{b}{k}\cos \mathrm{YP} = \frac{c}{l}\cos \mathrm{ZP}.$$

The calculations and observations were made for six important planes of the crystal with the following results:—

Plane.	Spacing.		Plane.	Spacing.	
	Calc.	Obs.		Calc.	Obs.
100	7·57	7·58	011	4·30	4·33
010	6·04	3·04	110	4·72	4·80
001	6·10	3·08	$\bar{1}01$	5·23	5·37

These results show that the spacings of the observed planes are all in order with the exception of d_{010} and d_{001}, the spectra of which show them to be, in fact, half the calculated values. This means, if we first place one structural unit (in this case, a molecule of tartaric acid) at each corner of the unit cell, the other structural unit must lie somewhere along the line OP (fig. 1) but not in such a way as to bisect OP, since d_{100} is normal. But the question immediately arises: Why are not d_{110} and d_{101} halved also? If the structural units lie at O and P and are similarly orientated to those lying at A, B, C, D, etc., then d_{110} and $d_{\bar{1}01}$ *must* be halved (though, as observed, d_{011} is unaltered).

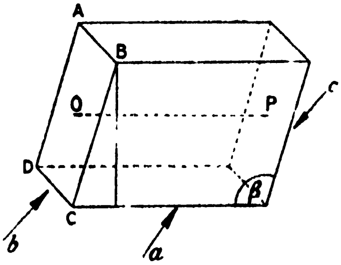


FIG. 1.

Hence we must conclude that either

(1) The units at O and P are differently orientated from those at A, B, C, D, etc., or

(2) They are displaced along the line OP at some distance not equal to OP/2, or

(3) Both (1) and (2) take place.

Such an arrangement would give the right number of molecules per cell and also agree with the above observed data.

As a matter of fact, we can see from symmetry considerations that, whether (2) is true or not, at least (1) must be true for an unsymmetrical structural unit; for the crystal structure of tartaric acid possesses but one element of symmetry, a unique dyad axis parallel to the crystallographic b axis, and the relative orientations of the two units in the cell must be such as to lead to the presence of this dyad axis of symmetry when the structure is completed. If the molecules at O and P are like those at A, B, C, D, etc., but rotated through 180° with respect to them, we can see at once how the axis of symmetry arises, and also, at the same time, account for the fact of the non-halving of d_{110} and $d_{\bar{1}01}$. It is interesting to note, as has been pointed out by G. Shearer, that if we assume that the structural units are unsymmetrical, we may, from purely theoretical considerations, arrive at the number and relative orientations of the structural units and limit their possible positions in the fundamental lattice. Consider a structural unit at the origin of a space lattice. Let it be represented by a system of points whose co-ordinates are

$$\begin{array}{ccc} x_1 & y_1 & z_1 \\ \cdot\ \cdot\ \cdot & \cdot\ \cdot\ \cdot & \cdot\ \cdot\ \cdot \\ \cdot\ \cdot\ \cdot & \cdot\ \cdot\ \cdot & \cdot\ \cdot\ \cdot \\ \cdot\ \cdot\ \cdot & \cdot\ \cdot\ \cdot & \cdot\ \cdot\ \cdot \\ x_n & y_n & z_n \end{array}$$

In the case under consideration, that of tartaric acid, the structural units are probably the molecules themselves, the shape and orientations of which may be defined by a system of points such as the centres of the atoms constituting the molecules. Suppose now the space lattice to be constructed of a set of units of a certain orientation, which will be, of course, similar for each molecule of the lattice, and which we may designate by the letter A. Then the co-ordinates of the system of A units may be written

$$\begin{array}{ccc} x_1+la, & y_1+mb, & z_1+nc. \\ \cdot\ \cdot\ \cdot & \cdot\ \cdot\ \cdot & \cdot\ \cdot\ \cdot \\ \cdot\ \cdot\ \cdot & \cdot\ \cdot\ \cdot & \cdot\ \cdot\ \cdot \\ \cdot\ \cdot\ \cdot & \cdot\ \cdot\ \cdot & \cdot\ \cdot\ \cdot \\ x_n+la, & y_n+mb, & z_n+nc. \end{array}$$

when l, m, n are integers, positive or negative.

Now rotation through 180° about the b axis combined with shifts αa, βb, γc parallel to the axes will give rise to a new set of units, a lattice of type B say. Its co-ordinates may be written

$$-x_1 - la + \alpha a, \qquad y_1 + mb + \beta b, \qquad -z_1 - nc + \gamma c.$$

. . . .

. . . .

$$-x_n - la + \alpha a, \qquad y_n + mb + \beta b, \qquad -z_n - nc + \gamma c.$$

Now for the complete structure to repeat through space, the same process applied to B units must necessarily reproduce A units. If we do this, the co-ordinates of the new A units are

$$x_1 + la, \qquad y_1 + mb + 2\beta b, \qquad z_1 + nc.$$

. . . .

. . . .

$$x_n + la, \qquad y_n + mb + 2\beta b, \qquad z_n + nc.$$

That is, for this new set to coincide with a set of A units,

$$2\beta b = 0, \text{ or } b \text{ (or } 2b, \text{ etc.)}, \textit{ i.e.}, \beta = 0, \text{ or } \tfrac{1}{2};$$

i.e. the B units must lie in the ac planes or half-way between them.

This reasoning, we see, leads us nearly as far as the experimental data given above. It presupposes the existence of an unsymmetrical unit. Whether the converse reasoning holds, that, given there are only two units in a fundamental cell of this class of symmetry, then the units are necessarily unsymmetrical, does not seem quite so clear, although it is very probably true. In fact, Shearer has suggested the principle that a crystal never uses more than the minimum number of units necessary to build up the lattice and at the same time comply with the symmetry conditions. Experiment has shown that this principle is verified in the case of tartaric acid and that, if we had known for certain that its molecules were quite unsymmetrical, we might have predicted, from purely crystallographic data, the number and orientations of the molecules in the unit cell and also limited their possible positions in the manner described above. The experimental work has shown just a little more from its first results, and now we are in a position to visualise, geometrically at least, the fundamental cell of the crystal lattice of tartaric acid. It is shown in fig. 2. The molecules at the corners of the cell are pointing in one direction while those lying along the central line point in a diametrically opposite direction. In the figure these latter are placed provisionally at the centres of the (100) faces, because the evidence given up to this point does not fix their position along the central line. It is clear that such an arrangement involves a unique dyad axis, parallel to the crystallographic b axis, for the completed structure, which will consist of two interpenetrating monoclinic lattices, either of which may be obtained from the other by a rotation of 180° about the *b* axis, combined with a certain shift to be defined later. In the completed structure, as in an ordinary crystal, this operation simply means congruence obtained after rotation about a unique dyad axis parallel to the crystallographic *b* axis. This is the only symmetry shown by crystals of dextro- or lævo-tartaric acids, and, therefore, up to the present, the deductions are in accordance with facts.

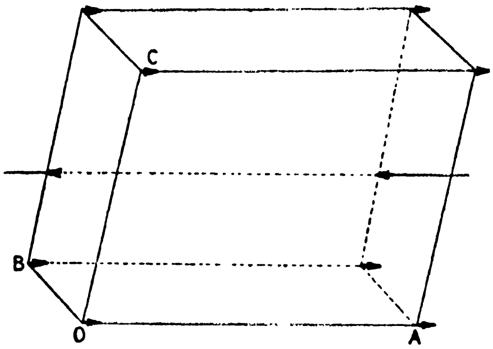


Fig. 2.

For further evidence, resort was made to intensity measurements of the reflections from the crystal faces previously mentioned. From the first, in view of the fairly complex formula which chemists have assigned to tartaric acid, it seemed very probable that not too much reliance could be placed on the evidence of intensity measurements. Even if these latter were taken with the greatest possible care, it is very doubtful if one is justified in comparing them rigidly with the theoretical values which might be calculated by using the simple Bragg theory. The complexity of the structure does not warrant this at the present stage of our knowledge concerning the scattering of X-rays. The labour of such calculations is very large, if precision is desired, and it is difficult to fix definitely the position of a small diffracting body in an elaborate structure. Rarely more than two theoretical intensity ratios can be made to agree with experiment at the same time. Considering these facts, the intensity measurements quoted below were made with care in the case of the first three planes only; the accuracy in the case of the other three is only very moderate.

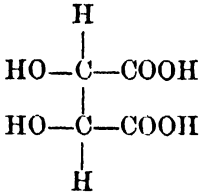
Approximate Relative Intensities in Arbitrary Units.

Plane.	First order.	Second order.	Third order.	Fourth order.
100	29	20	38	$2\frac{1}{2}$
010	99	11	1	–
001	61	15	2	—
011	178	2	6	3
110	27	3	8	1
$\bar{1}01$	54	15	18	9

The outstanding features of this Table are the large third order reflection shown by the (100) face and the enormous value of the first order reflection from the (011) face. All the faces examined proved extremely good reflectors, but the first order reflection from (011) was so large as to stand in a class by itself. It is to be noticed that the second order reflection from the same face is relatively insignificant.

At this stage we cannot make much further use of these measurements, and so it would be well now to look a little deeper into the question of the probable atomic structure of the crystal.

The accepted structural formula for tartaric acid is



$$\begin{array}{c} H \\ | \\ HO—C—COOH \\ | \\ HO—C—COOH \\ | \\ H \end{array}$$

when written all in one plane. It contains two "asymmetric" carbon atoms and exists in two optically active forms (*d*- and *l*-tartaric acids), one "internally compensated" form (mesotartaric acid), and one "externally compensated" form (racemic acid). This paper is concerned with the structure of the dextro- and lævo-forms only. The existence of these four forms has been satisfactorily explained by chemists by Le Bel and van't Hoff's theory of stereoisomerism. This theory ascribes the existence of the four forms to the presence within the molecule of two asymmetric carbon atoms which may be considered to be situated at opposite ends of a dumb-bell, so to speak. Around these in space are arranged the other six groups (—OH —H and —COOH round each asymmetric carbon atom), in such a way that their relative orientations are the deciding factors as to whether the structure is optically active or not. The usual way of representing the possible arrangements is shown in fig. 3, in which the asymmetric carbon atoms are placed at the centres of regular tetrahedra.

Forms (*a*) and (*b*) represent the two optically active acids, while (*c*) represents the "internally compensated" meso-tartaric acid. The "externally compensated" acid (racemic) is a mixture of equal proportions of (*a*) and (*b*). It is to be noticed that in either active modification the order of sequence of the —OH, —H and —COOH groups is the same round both asymmetric carbon atoms of the molecule, but in the meso-form the order in which these groups succeed each other round one of the asymmetric carbon atoms is opposite to the order in which they succeed each other round the other asymmetric carbon atom. In other words, if we look along the bond joining the two asymmetric carbon atoms in the direction from the apex of the tetrahedron to a carbon atom, we see the three subsidiary groups succeeding each other either clockwise or anti-clockwise according as we are examining one active modification or the other; but in the inactive modification we see a clockwise sequence as we look towards one carbon atom but an anti-clockwise sequence when we look towards the other carbon atom. This is all that Le Bel and van't Hoff's theory stipulates. The present examination of the crystal structure of tartaric acid not only tests, it is believed, the validity of the theory, but also provides evidence as to the detailed orientations of the —OH, —H and —COOH groups and helps to decide actually the structural difference between the dextro- and the lævo-acid.

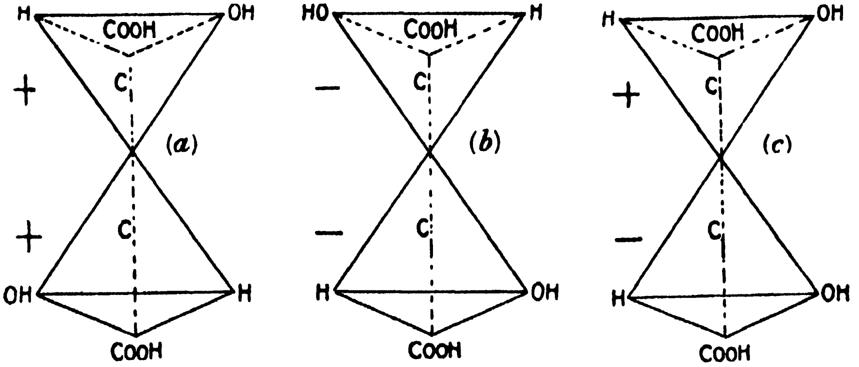
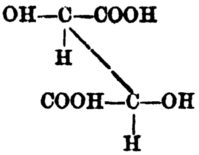


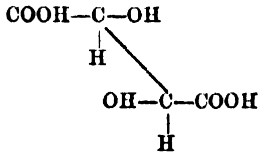
Fig. 3.

Corresponding to the two optically active forms are the two enantiomorphous forms of the monoclinic sphenoidal class of symmetry to which tartaric acid crystals belong. The chemical diagrammatic formulæ show well this mirror-image property of the two active modifications, and it would be well, before proceeding further, to set out diagrammatically the possible enantiomorphous configurations of the constituent groups of the molecule.

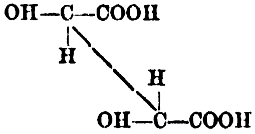
I.



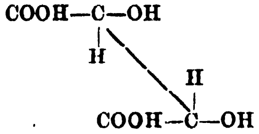
and



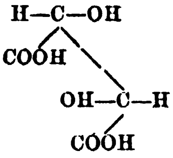
II.



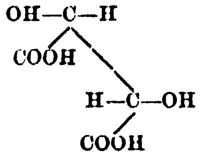
and



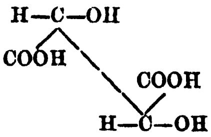
III.



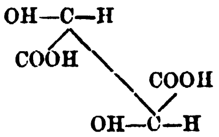
and



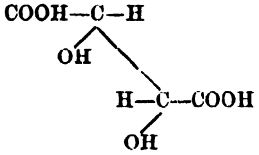
IV.



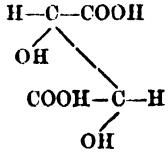
and



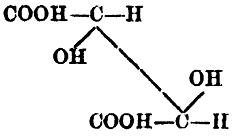
V.



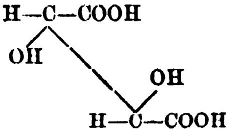
and



VI.



and

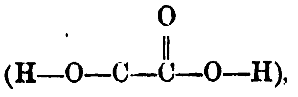


Now it is a significant fact that, if we accept W. L. Bragg's values of the "atomic diameters" (1·30 A.U. for oxygen and 1·54 A.U. for carbon), and also make use of the fact observed by W. H. Bragg that the "atomic diameter" of hydrogen in naphthalene is approximately 1 A.U., the length of the *a*-axis of tartaric acid is almost exactly equal to twice the length of a (COH) group. Thus

$$2(1{\cdot}54+1{\cdot}30+1{\cdot}0) = 7{\cdot}68 \text{ A.U.}$$

Whereas, as shown above, $a = 7{\cdot}693$ A.U.

It is difficult to believe that such an equality is a mere coincidence; we must adopt the configurations (I) and (II) given in the list above, for writing (OH—C—COOH) in the more detailed form



$$(\mathrm{H{-}O{-}C{-}\overset{\overset{O}{\|}}{C}{-}O{-}H}),$$

it is clear that if the molecules lie lengthwise along the *a*-axis of the lattice, they will readily fit into the structure. Again, if we adopt this suggestion, remembering at the same time that the complete structure holds together in virtue of the fact that it consists of two interpenetrating monoclinic lattices, either of which may be obtained from the other by a rotation through 180°, combined with a certain shift, it seems probable that the two asymmetric carbon atoms of each molecule will lie, more or less exactly, along the unique dyad axis of the crystal. In other words, we may consider the dumb-bell nucleus of the molecule to bear some close relation to, if it does not actually coincide with, the axis of rotation, which is the essential factor connecting the two interpenetrating lattices. Thus, taking these dumb-bell pairs of carbon atoms as the corresponding points of the fundamental lattice, we arrive at such a basis for our structure as is represented in fig. 4. For the present, we leave the exact position of the pairs along the line OP open. We are left now with the ketonic oxygen (=O) of the —COOH group and the remaining hydrogen atom of the asymmetric carbon atom to fix in such positions that they may be of use in the process of connecting up the molecules in the cell. Consider now the face ABCD (figs. 4 and 5).

$$MN = c/2 = 3{\cdot}1 \text{ and } LM = b/2 - 1{\cdot}54 = 3{\cdot}02 - 1{\cdot}54 = 1{\cdot}48.$$

$$\therefore\ LN = 3{\cdot}43, \textit{ i.e.}, QR = 1{\cdot}89.$$

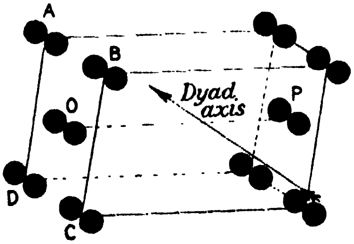


Fig. 4.

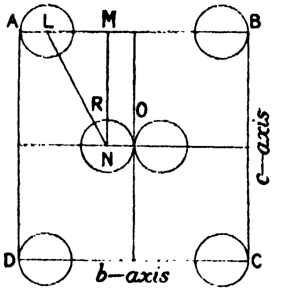


Fig. 5.

Thus, as the arrangement now stands, there is not quite enough room for an (OH) group to lie along QR between the molecules A and O. But the displacement of O and P along the direction OP is not yet fixed; in fact, it would require very little displacement of O and P to make room for (OH) groups to act as supports between the two interpenetrating lattices. Furthermore, we can see at once how it is possible, by means of these (OH) groups, to link up the whole structure in an extremely simple manner (fig. 6), if we grant that the main body of the molecule lies approximately parallel to the long axis of the cell. The diagram shows how it becomes necessary to adopt the configuration (II) given above, and also shows clearly how the structure holds together in virtue of the fact that either of the two interpenetrating lattices is rotated through 180° with respect to the other. It is to be particularly noticed, too, that no new element of symmetry has been added by the process. This, of course, is as it should be, since we started out with the conception of an unsymmetrical structural unit.

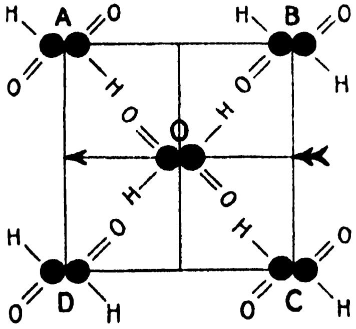


FIG. 6.

Another point must be made clear: that we are provisionally assuming the molecule to have the flat appearance of the chemical formula. Since fig. 6 is only a projection on to a plane perpendicular to the long a-axis, each dumb-bell pair of carbon atoms, shown in accordance with the provisional assumption, hides another dumb-bell pair, the ketonic oxygens (=O) being connected to the carbon atoms of the (—COOH) groups and the hydrogen atoms (—H) to the carbons of the asymmetric doublet. The point is made clear by another aspect of the cell (fig. 7), in which the asymmetric carbon atoms are shown with rings round them. To avoid confusion, one linking only is shown at each end of the cell, but the positions of the others may be readily found with the aid of fig. 6. Fig. 8 shows a projection of the same

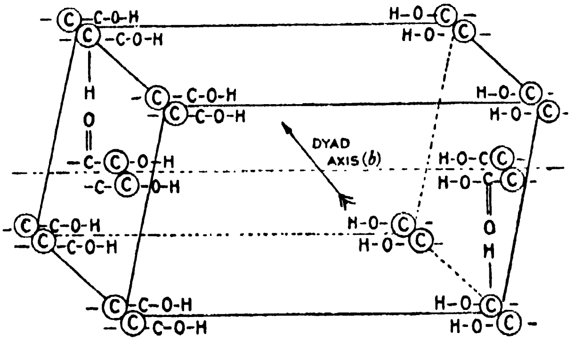


FIG. 7.

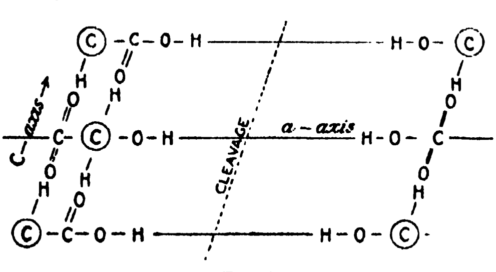


FIG. 8.

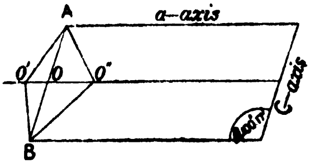


FIG. 9.

on the *b*-plane (010). Fig. 8 shows also an objection to the arrangement as it now stands. There are two hydrogens (—H) and two ketonic oxygens (=O) in each molecule, and the bonds between each of these and their respective carbon atoms make certain angles with the bonds connecting the asymmetric and ordinary carbon atoms. In fig. 8 the former class of bonds are drawn parallel to the *c*-axis, and the latter class parallel to the *a*-axis, and such a representation obviously means that one (C—H) bond makes a certain angle with the (C—ⓒ) bond, while the other makes the supplement of this angle; and similarly for the angles between the (C=O) bonds and the (C—ⓒ) bonds.

There seems to be no reason whatsoever why this should be so; in fact, so far as we know from chemical considerations, one half of the ordinary tartaric molecule behaves exactly like the other half and is indistinguishable from it. The only argument in favour of the above diagrammatic representation is that it seems (until further evidence is considered) the only relative orientation which makes the gaps for the (=OH—) linkages all equal. Since the lattice is monoclinic, displacement of the molecule *o* along the *a*-axis only makes matters worse; in positions such as o' or o'', not only the angles are unequal but the (=OH—) gaps are unequal also. And such a possibility as the latter cannot be admitted. Obviously then, if the line of reasoning we are now following is the right one, there is still some further displacement and distortion possible and necessary before such objections as the above (and others) can be met satisfactorily. We shall see that consideration of the intensity measurements leads to the solution of the problem.

The theoretical relative intensities have been worked out by aid of the polygon of amplitudes and phases, assuming the Bragg hypothesis that the amplitude of a diffracted wavelet is proportional to the weight of the diffracting atom. With such a complex structure as the one under consideration it is laborious to arrive at precise results, but sufficient has been worked out to show the main lines of argument to be adopted. We will consider first the intensities from the (100) plane. In fig. 8 this plane is shown perpendicular to the plane of the paper. Then, calling the arrangement shown there the "zero" arrangement, or the one in which the relative displacement of the two lattices is zero (reckoning positive displacements of the middle lattice to the left) the following values have been arrived at:—

Displacement.	Amplitudes.			Intensities.		
	1st order.	2nd order.	3rd order.	1st order.	2nd order.	3rd order.
−0·75 A.U.	44	33	35	1936	1089	1225
0·0 „	49·5	29	18	2450	841	324
+0·75 „	73	80	31	5329	6400	961
+1·50 „	65	43	95	4225	1849	9025

The theoretical values of the intensities given in the last column are uncorrected for any law of falling-off of intensity of reflection with order of reflection. The experimental intensities for the (100) plane are approximately in the ratio 3 : 2 : 4; that is, if we assume an inverse square law for the falling off of intensity with order of reflection (the law which has been found to hold approximately when the diffracting atoms lie on equidistant and similar planes) the experimental values should correspond to theoretical values 3 : 8 : 36.

Now it is easily shown that such theoretical values are impossible in the case of the similar and equi-distant planes referred to, for the summation for two equal vectors gives a maximum third order when their phase difference is 120°. The ratios are then 1 : 1 : 4. Further increase of phase-difference decreases the first order, but brings the second and third orders nearer equality. In view of these facts it is justifiable to aim only at such an arrangement as will give a large third order reflection and, if possible, at the same time show some sort of approximation to theoretical values of the ratios for the different orders of reflection. In our case, the third order attains its maximum value when the displacement of the middle lattice (fig. 8) is about 1·5 A.U. to the left, and in this position the ratio for the second and third orders is about right too. This arrangement has accordingly been adopted. It corresponds to a rotation of the molecules about the line joining the centre of the two asymmetric carbon atoms (fig. 10). As is seen later, there is some possibility that a displacement of rather more than 1·5 A.U. would be necessary to make room for the (=OH—) linkages, but as the additional displacement cannot be very large and cannot be decided definitely, the argument is unaffected and we may proceed to the consideration of the next point.

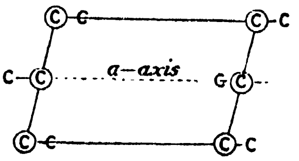
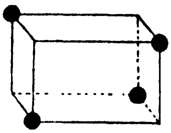


Fig. 10.

Fig. 5 reveals a strong objection to the present state of development of our structure. It shows that the system now consists of almost flat layers of atoms parallel to the plane (001). It is *a priori* improbable that such is the case, for it would mean that the majority of the "bonds" between the atoms would lie parallel and in one plane. But a more powerful argument against such an arrangement is that it involves almost complete halving of d_{010}. Fig. 5 shows the point well, LM = 1·48 = 3·02/2 nearly. And if the (010) planes were interleaved half-way by almost identical planes, that would mean a large second order and a negligible first order for these planes. Experiment shows that such is not the case, in fact, the observed first order intensity is nine times the second order intensity. Now any departure from the planar arrangement so far arrived at would help to bring about an agreement between the calculated and observed intensities from the plane (010), and when we consider the arrangement of carbon atoms in diamond, graphite, and the organic compounds, the structures of which Sir William Bragg has recently elucidated, it seems very probable that some sort of tetrahedral arrangement of the four carbon atoms and their "bonds" is much more likely to be at the root of such a structure as the one under consideration. We shall see that the tetrahedral arrangement affords a very satisfactory interpretation of the observed facts.



In the first place, we notice that a tetrahedral arrangement such as in bringing the carbon atoms along the *b*-axis nearer together, will help to move to a new position the plane which has hitherto been halving d_{010} with an almost equal weight. The observed ratio for the first and second order intensities from the plane (010) is 9 : 1, and if we multiply this by 4 to allow for a "diminution with order," the calculated intensities for (010) should be as 9 : 4; that is, the amplitude ratio should be 3 : 2. We may now calculate approximately the phase difference necessary to give this ratio. The ratio of the amplitudes of the first and second order reflections from two equal diffracting centres, whose difference of phase is α, can be shown to be $\cos\frac{1}{2}\alpha/\cos\alpha$, *i.e.*, $\cos\frac{1}{2}\alpha/\cos\alpha = \frac{3}{2}$, whence $\alpha = 112^\circ$ approximately. But $d_{010} = 3{\cdot}0$, therefore the dividing planes must be at a distance $(112 \times 3)/360 = 0{\cdot}93$ from the principal planes. Then, if we take the diameter of the carbon atom to be 1·5, it also follows that the spacing d_{001} (3·1) is interleaved at distances about 1·16, from which we see that the second order intensity from the (001) plane should be more nearly equal to its first order intensity than is the case with the (010) plane. This agrees with observation. If we now test the (011) plane, we find that the calculated (uncorrected) intensities are approximately in the ratio 314 : 17, which agrees very well with experiment. In fact, on the whole, it is clear that the proposed tetrahedral arrangement explains the main features of the intensity measurements, since exact agreement with theory is not to be expected. But now a more interesting fact is revealed. The proposed tetrahedral arrangement has automatically put us in a position to make the further distortion which is necessary to equalise the (=OH—) linkages and the angles which they make with the rest of the molecule.

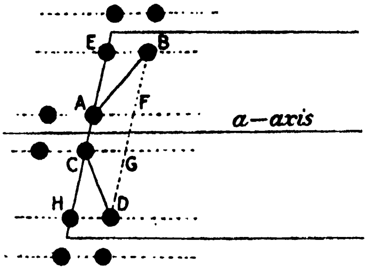


Fig. 11.

At present, our carbon atoms are situated at alternate corners of a rectangular parallelepiped. Fitting these into the monoclinic lattice and projecting on to the (010) plane gives fig. 11. This shows the linkages AB and CD still unequal and still making unequal angles with the *a*-axis. But it is now apparent that, if the parallelograms AEBF and CGDH were also rectangles, then AB would be equal to CD, and also $\widehat{BAF}$ would equal $\widehat{DCG}$. This means that, to make the structure consistent, we have to make the carbon parallelepiped more oblique, as shown in fig. 12. The angle ACG now becomes approximately equal to $\tan^{-1}\dfrac{2\times 1\cdot 16}{c\cos(180-\beta)}\approx 64°$, and the nucleus of carbon atoms can in future be represented by the oblique parallelepiped (fig. 13), in which A and B represent the asymmetric carbon atoms and C and D the carbon atoms of the (—COOH) groups. The bonds must lie along the directions AB, AC, BD, where AB $\approx$ AC $\approx$ BD $\approx$ 1·5 A.U. and $\theta \approx 64°$.

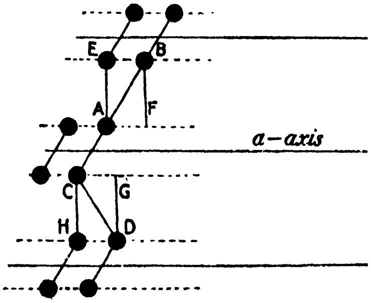


Fig. 12.

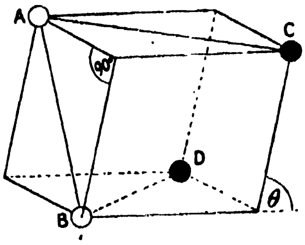


Fig. 13.

Fig. 14 (*a*) shows completed one of the edges of the unit cell parallel to the *a*-axis. Fig. 14 (*b*) is an end view of the same, which indicates also the approximate positions of the linkages.

It is to be noticed that it has become necessary for the (—OH) groups to link across diagonally, as shown in the figure. This is a very satisfactory conclusion, for the arrangement with which we started out, namely, long strings of atoms parallel to the *a*-axis, seemed to be "too good to be true," so to speak. At the same time, this diagonal arrangement of (—OH) groups involves another conclusion, that the distance between the centres of oxygen and hydrogen in an (—OH) group is certainly more than 1·15 A.U. Sir William Bragg's investigation of the crystal structure of ice* leads one to expect this. In ice, the distance between the centres of oxygen and hydrogen is 1·38.

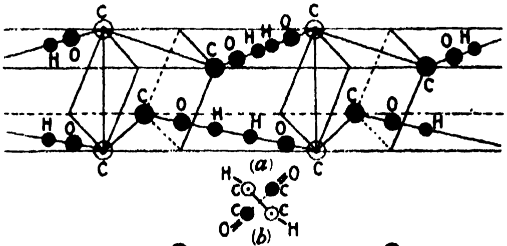


Fig. 14.

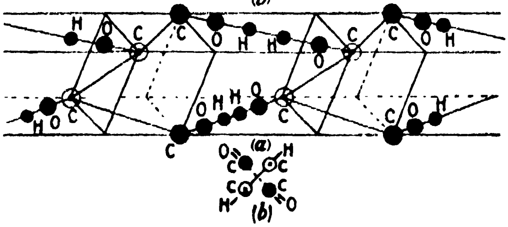


Fig. 15.

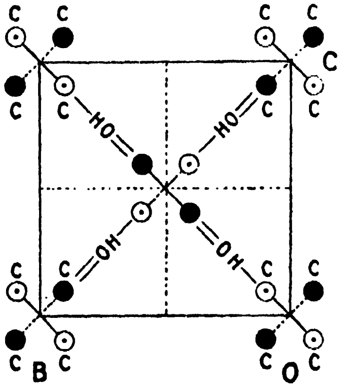


Fig. 16.

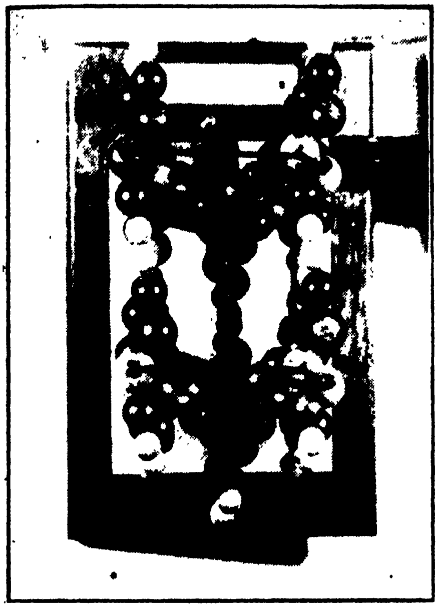
Another point worthy of notice is that the length of the (=OH—) linkages, which contain a ketonic oxygen atom (=O) must be rather less than the length of an ordinary (—OH) group. This also is to be expected.

The enantiomorphous form of fig. 14 is shown in fig. 15. Fig. 16 shows a complete end view of the unit cell corresponding to fig. 14. Fig. 14 is a diagram of the edge parallel to the *a*-axis of the unit cell, whose (100) face is shown in fig. 16. Three photographs are also given of the complete model of the tartaric acid structure, which has been constructed on the scale of 1 cm. to 1 A.U. (p. 521).

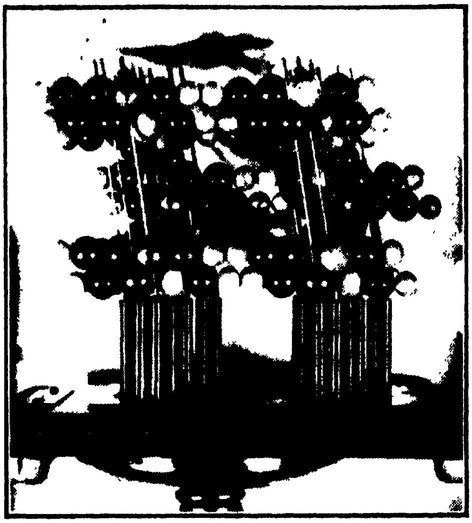
---

* 'Proc. Phys. Soc.,' vol. 34 (April, 1922).

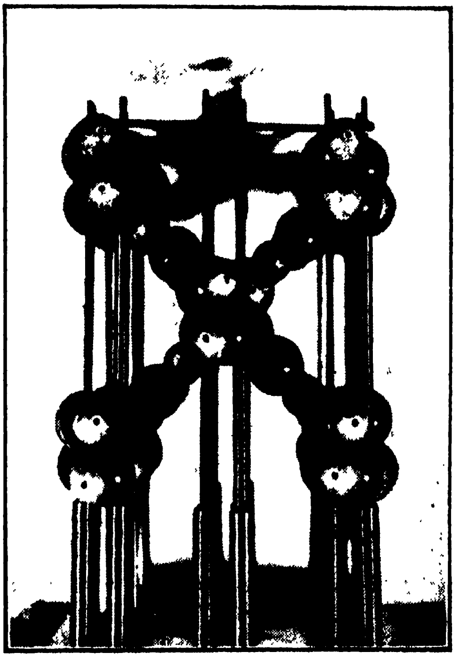
MODEL OF TARTARIC ACID STRUCTURE.



Face (001).



Face (010).



Face (100).

2. *Properties.*

The crystalline structure arrived at in the previous section explains in a very satisfactory manner the characteristic properties of active tartaric acid. Not only does it seem to be correct from the purely crystallographic considerations so far mentioned, but it also shows other interesting features, which appear to throw light on certain chemical and physical problems.

The crystals show one perfect cleavage only, parallel to the plane (100). This fact the crystal structure brings out very clearly, for it possesses only one plane parallel to which we should expect a perfect cleavage. This is the plane about which lie the junctions of the hydroxyl groups. Figs. 7, 8, 14 and 15 illustrate how the molecules are held together, end to end, by forces between the hydrogen atoms of adjacent hydroxyl groups, one hydroxyl group being carboxylic and the other alcoholic. The plane about which these molecular junctions lie is the plane (100), the perfect cleavage plane of the crystal. There are none of the strong valency forces acting across it, only the weaker forces which act from molecule to molecule and serve to hold them together in the crystal lattice. It is a remarkable fact that similar cleavage planes are clearly shown in the organic crystals investigated by Sir William Bragg. In naphthalene, for instance, the perfect cleavage passes across the junctions between the β-hydrogens of adjacent molecules.* Tartaric acid thus affords another example of that fundamental difference between organic crystals and those of the inorganic polar type, such as sodium chloride. In the latter the identity of the molecule is lost in organised ionisation, but in organic crystals each molecule seems to preserve its individuality.

The junctions in the (=OH—) linkages seems to be of a different type from the junctions between the hydrogens of the hydroxyl groups. They are much stronger, for no cleavage is observed across them. This is undoubtedly to be associated with the fact that the oxygen is a "double-bonded" oxygen in this case, and therefore in common with most other ketonic oxygens known to organic chemistry, has a marked tendency to seize upon and attach itself to other available groups. It is interesting to note that such a tendency can also be utilised for the purposes of crystal architecture. The distance between the centres of carbon and oxygen in the (—C=O) group seems to be rather less than the distance between corresponding atoms in the (—C—O—H) group, a fact which is rather to be expected from the theory of electron-sharing.

The reason for the intense first-order reflection from the (011) plane is made clear by an inspection of fig. 16, or the (100) face of the model. The density of atoms per unit area is greater for the (011) plane than for any other plane of the lattice. It is this circumstance, aided by the interleaving which tends to throw most of the energy into the first-order spectrum, that gives rise to that powerful reflection which distinguishes the (011) plane from all the other planes. It is also a noteworthy fact that the (011) face is often the best developed on the crystals, which is again what the model would lead one to expect, when we remember that a crystal tends to develop only those planes of greatest reticular density.

---

* 'Proc. Phys. Soc.,' vol. 34 (December, 1921).

The two enantiomorphous forms are shown diagrammatically in figs. 14 and 15. Either is the reflection of the other across the plane (010) (the symmetry plane of the monoclinic system). Good crystals of the lævo-acid were not available, and thus it has not been possible up to the present to test for possible differences with the ionisation-spectrometer. However, it may be shown from theoretical considerations that no difference whatsoever is to be expected. It may be stated, as a general proposition, that it is impossible by the diffraction of X-rays to detect directly the difference between a dextro- and a lævo-compound. For the only difference between two enantiomorphs is embodied in the fact that either is the mirror image of the other across the symmetry plane. To every face above the plane of symmetry (the plane of the paper in a stereographic projection) there is a corresponding face below, whose spacing is the same and whose molecular and atomic orientations are the mirror image of the molecular and atomic orientations above the plane. The only difference between a right- and a left-handed form is that the spacings are interleaved in the opposite sense, thus:—

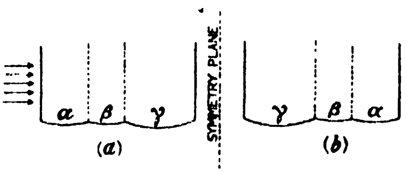


FIG. 17.

In the figure, (*a*) is the mirror image of (*b*) and *vice versâ*, and therefore since the diffraction of X-rays is a summation effect of all the planes concerned, there can be no difference between the diffraction patterns produced by the two enantiomorphous forms. If we could by means of the diffraction figure detect the difference between a dextro- and a lævo-body, we could detect the difference in diffraction pattern produced from the opposite sides of a crystal face. The case of tartaric acid may be illustrated by a section perpendicular to the a-axis (fig. 18).

It is clear that the back of one form is, with respect to incident rays, similar to the front of the other, and *vice versâ*.

The optical activity of tartaric acid may be observed either in the crystalline form or in solution. The optical activity is thus a property of the molecule itself. Indeed, it is not to be expected that a unique dyad axis would give rise to optical activity in a crystalline structure, for a clockwise rotation through 180° is indistinguishable from an anti-clockwise rotation through the same angle. In quartz the rotation is a property of the crystal itself,* and is associated with a spiral arrangement of the silicon atoms about the triad axis. If we are to explain the rotatory power of the tartaric acid molecule in a similar manner, it is natural to look into the structure we have arrived at for some spiral arrangement to which the optical activity may possibly be ascribed. When we examine the model of the structure, we are at once struck by the presence of the obvious spiral arrangement shown in figs. 14 and 15. From independent considerations we have concluded that the four carbon atoms which form the nucleus of the molecule are situated at alternate corners of an oblique parallelepiped, and by adopting such an arrangement we have automatically produced an irregular spiral formation within the molecule itself. Again, this procedure, as we saw above, has forced us to introduce another spiral arrangement of hydroxyl groups, for we cannot effect a junction between the hydrogen atoms unless we reverse the twist which is associated with the nucleus of the molecule. Thus, at first sight, we seem to have found more than sufficient structural material to which we may ascribe the optical rotation, but we shall see below that this doubly twisted arrangement has probably a very important physical significance. For the present, it is to be noticed that the twist in one enantiomorph is in the opposite sense to the twist in the other, as is to be expected, and also, as we see on examining the model, this reversal of twist involves a reversal of the order in which the (—H), (—OH), and (—COOH) groups succeed each other round the asym-

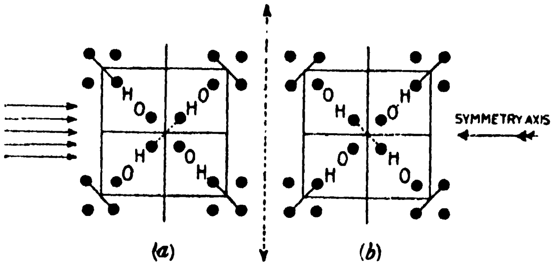


FIG. 18.

---

* Bragg, 'X-rays and Crystal Structure,' p. 166, 1st edition.

metric carbon atoms. This, too, is to be expected, though hitherto the chemical theory of stereo-isomerism has not indicated any definite physical connection between the order of sequence of these groups and the optical activity of the molecule. It appears now that the rotation is to be ascribed primarily to the nucleus of the molecule. The reversal of optical rotation necessarily involves a reversal of the order of sequence of the subsidiary groups, in a way that consideration of the crystal structure makes apparent at once.

Another point worthy of notice is that the irregular spiral formations to which we have ascribed the rotatory power of the crystal possess similar symmetry to that shown by the complete lattice. They exhibit in themselves the unique dyad axis of the structure. In tartaric acid, we know that the plane of the optic-axis is perpendicular to the symmetry plane of the monoclinic system, the obtuse bisectrix coinciding with the dyad axis; that is, the optic axes make equal angles with the crystallographic axis and the rotatory power along either of them is the same ($8\frac{1}{2}°$ per millimetre for red light). This symmetrical distribution of optical activity is in agreement with our hypothesis concerning its physical basis, for though the tartaric molecule as a whole is completely devoid of symmetry, yet the parts of it to which we have ascribed the rotatory power possess the symmetry corresponding to the known distribution of this rotatory power.

Biot showed that solutions of ordinary (dextro-) tartaric acid in water, alcohol, or wood-spirit do not accord with his law of rotatory dispersion ($[\alpha] \propto 1/\lambda^2$). On the contrary, they show a maximum rotation in the green, a special property which combination with bases seemed to destroy. He also showed that, approximately, the specific rotation of an aqueous solution at a given temperature is a linear function of the concentration;

or
$$[\alpha] = A + B\epsilon,$$

where ϵ is the proportion of water in the solution. This formula is not applicable to very dilute solutions. He then extrapolated to obtain the rotation of the anhydrous acid and predicted that for red light this would change sign at 23° C., a prophecy which was verified some years later.*

The addition of boric acid to tartaric acid solutions greatly increases the dextro-rotatory power. Heating or dilution of a tartaric solution displaces the maximum towards the violet end of the spectrum and increases the rotation. Supersaturated solutions give a negative rotation, becoming positive for longer wave-lengths. To account for this peculiar behaviour of tartaric acid, Arndtsen in 1858 put forward the following hypothesis:—

---

* Prof. Lowry has given a very complete review of the subject of the rotatory dispersion of tartaric acid and the tartrates in his Bakerian Lecture of June, 1921. In the preparation of this paper full use has been made both of his summary of previous work and of the account of his own work on the subject. But in this paper it is proposed to put forward an alternative hypothesis to the one which was advanced in his lecture to explain the anomalous rotatory dispersion of the tartrates. A short restatement of his facts is necessary.

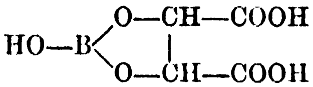
"If one should imagine two active substances which do not act chemically on one another, of which one turns the plane of polarisation to the right, the other to the left, and, in addition, that the rotation of the first increased (with the refrangibility of the light) more rapidly than that of the other, it is clear that, on mixing these substances in certain proportions, one would have combinations similar to those of tartaric acid."

Lowry suggests that these two opposing bodies may be the original acid and either an ion, a hydrate, a polymer or an isomer, and proceeds to rule out the first two on the ground of the anomalous behaviour of the non-crystalline acid and the liquid esters. On the same grounds, he also rules out any molecular rearrangement which involves the hydrogens of the carboxyl groups.

Biot stated that the alkali tartrates are normal in their rotatory dispersion, but Lowry has shown that they are slightly anomalous with a large positive and a small negative term. To tartaric acid itself he ascribes the Drude formula

$$\alpha = k_1/(\lambda^2-\lambda_1^2)-k_2/(\lambda^2-\lambda_2^2),$$

an equation which indicates the presence of two opposing rotatory systems. But the boric acid derivative, of which the probable formula is



is fixed and "simple" in its rotatory dispersion. Similarly, tartar-emetic has also a large dextro-rotatory power which is of the "simple" kind.

Lowry suggests the existence of "dynamic isomerism" to explain the above facts. He argues that there must be existing side by side two opposing isomeric forms of tartaric acid which are in dynamic equilibrium. He has not, however, suggested the two possible chemical formulæ for these two dynamic isomers, neither have they yet been isolated separately.

The crystalline structure of tartaric acid seems to afford a simple explanation of all the facts. The fundamental cell is constructed of molecules which show *two* spiral arrangements, one associated with the four carbon atoms forming the nucleus of each molecule and the other associated with the four hydroxyl groups. And, moreover, as reference to figs. 14 and 15 and the model shows clearly, the twists of these two systems are in opposite senses. It is now proposed that in these two spiral formations are located the two opposing rotatory systems which are sufficient to explain the anomalous behaviour of the acid and some of its derivatives. The study of the rotation of these compounds points to the fact that almost certainly the most stable system is the dextro-system which predominates in dilute solutions. In all ordinary circumstances "dextro"-tartaric acid and its derivatives are dextro-rotatory, but in the crystalline form and in certain non-crystalline states and supersaturated solutions the acid shows a lævo-rotation. (Dufet* has observed and measured the rotatory power of tartaric acid along its optic axes and states that its rotatory dispersion resembles that of quartz.) Decrease of concentration of tartaric acid solutions or rising temperature enhances the dextro-effect. "Fixing" the (—OH) groups by means of boric acid also greatly increases the dextro-rotatory power, etc. In short, the detailed consideration of this group of related compounds (which will be given later) leads one to the conclusion that the two extremes of rotatory power are the dextro in the very dilute solutions or when the hydroxyl groups are "fixed," and the lævo in the crystalline form. Through all the changes enumerated, there is only one part of the molecule to which we can ascribe any degree of stability, and that is the nucleus of four carbon atoms arranged in an irregular spiral. Such effects as ionisation and hydration would tend to destroy the lævo-rotatory action of the four hydroxyl groups, and it is probable that it requires the application of the forces which bind the molecules into the crystalline structure to complete the lævo-rotatory system. Any change, such as fusion, solution and certain chemical reactions, which leads to departure from the orientations which hold in the crystalline structure, will tend to eliminate the hydroxyl spirals. But throughout all these changes, there seems to be no reason why the arrangement of the nucleus of the four carbon atoms should be materially affected. *The balance of evidence leads to the conclusion that the dextro-rotatory property of ordinary tartaric acid is associated with the carbon nucleus of the molecule.*

The perfect cleavage of the crystals is a manifestation of the fact that the structure is ruptured across the junctions of the hydrogen atoms of the hydroxyl groups. In the complete structure the hydroxyl groups are linked together in such a way as to form a twisted arrangement of atoms. The act of solution or any departure from the crystalline form is sufficient to destroy the twisted arrangement. But the fact of crystal growth itself points to an initial leaning on the part of the hydroxyl groups towards such an arrangement as holds in the crystalline form, and it is fairly safe to assume that the tendency of the molecules to build themselves up into a lattice exists even in dilute solutions. There will always be present strings of molecules, connected together at the hydrogen junctions, which will show to a certain extent those properties which are completely manifested only in the crystalline state. On dilution, ionisation and hydration will break up these strings more and more, and the dextro-rotatory effect of the nucleus will increasingly predominate.

---

* Dufet, 'Journ. de Phys.,' vol. 4 (1904).

Other cases involved in this hypothesis will be considered in rather more detail later, when it is hoped to give an account of further study of the tartaric acids. The racemic acids are now being investigated. An account will also be given of an attempt to discriminate between the two enantiomorphs of active tartaric acid, by aid of a comparison with the crystal structure of quartz and of the analogy afforded by Reusch's combinations of mica plates. It will thus be shown that it is probable that in the dextro-acid the order of sequence of the (—H), (—OH) and (—COOH) groups is anticlockwise, as we look towards the asymmetric carbon atom in the direction leading towards it from its companion asymmetric carbon atom.

The aspect of the molecule as a whole shows a distorted tetrahedral arrangement of bonds.

Summary.

1. From observations made with an X-ray spectrometer the crystal structure of active tartaric acid is obtained.

2. The crystal structure is compared with known chemical and physical properties, and is found in the main to be in accordance with facts.

3. Le Bel and van't Hoff's theory of stereoisomerism is in its essentials confirmed, and the direct link between the crystallographic enantiomorphs and the chemical stereoisomers is revealed.

4. The rotatory properties of the acid are discussed with reference to the symmetry of the crystal and the spiral arrangement of atoms within the molecule.

5. It is shown that it is impossible by the diffraction of X-rays to distinguish between the dextro- and lævo-forms of an optically active body.

6. The model affords a simple explanation of the anomalous rotatory dispersion of tartaric acid and many of its derivatives. The dextro-rotatory power of ordinary tartaric acid is associated with the spiral arrangement of the four carbon atoms of the molecule.

In conclusion, I wish to thank Sir William Bragg for his kindness in allowing me to make use of the facilities of the Physics Research Laboratory at University College, and for his helpful criticism and the interest he has shown throughout the progress of the research. I wish also to thank Prof. A. W. Porter for his kindly advice and criticism, and Dr. O. L. Brady for supplying me with crystals and information on certain chemical points.

On the Quantum Theory of the Simple Zeeman Effect.

By A. M. MOSHARRAFA, B.Sc., University of London, King's College.

(Communicated by Prof. O. W. Richardson, F.R.S. Received September 1, 1922.)

§ 1.

The aim of this paper is to put forward a theory of the simple Zeeman effect which possesses the same general features as those of the corresponding theory in the case of the Stark effect already developed by Epstein and Schwarzschild. These general features may be briefly described as follows:—(1) The steady states of the atom are governed by classical dynamics subject to certain conditions, the quantum restrictions, which define the atom *both in the absence and in the presence of the field*; (2) radiation occurs during the period of transition from one steady state to another according to Bohr's energy relation $h\nu = W_m - W_n$.

The essential difference between our treatment and the already existing Bohr-Sommerfeld theory* lies in this: That the relation between the motion of the atom, in the presence and in the absence of the magnetic field respectively, as defined by classical dynamics, is not assumed in the former, whereas it plays a fundamental *rôle* in the latter. This relation (which we shall refer to as Schott's theorem)† involves for its classical proof a consideration of the motion of the atom during the period of establishment of the magnetic field. But the atomic system during that period is not a conservative one, and hence cannot be claimed to be fully comprehended from the point of view of the quantum theory. Thus it would seem desirable to avoid assuming Schott's theorem in a quantum treatment of the Zeeman effect. Our treatment, in which such an assumption is avoided, is, however, not incompatible with that theorem; in fact, it could be proved‡ that the latter follows as a necessary consequence of our analysis.

With regard to the quantum restrictions which define the steady states, we find it necessary to employ not the original form proposed for them by Dr. Wilson,§ but a slightly extended form first suggested by Prof. Wilson himself in the course of a discussion on the quantum theory at the Edinburgh meeting of the British Association last summer (1921). This latter form, represented by equations (17A) below, was put forward by Prof. Wilson on perfectly general grounds and apart from a theory of any particular phenomenon. The present writer was led to a special application of the same extension, represented by equations (17B) below, from a theoretical study of the Zeeman effect.

---

* See, *e.g.*, A. Sommerfeld, 'Atombau und Spectrallinien,' 2 Auf., 6 Kap, § 5, Braunschweig, 1921.

† See G. A. Schott, 'Electromagnetic Radiation,' Cambridge University Press, § 302, p. 317 (1912).

‡ See Appendix.

§ William Wilson, 'Phil. Mag.,' vol. 29, p. 796 (1915), set of equations (2).

We propose, in the first place, to show that if the original form of the quantum restrictions is maintained, *both in the presence as in the absence of the field*, then there can strictly be no difference between the energies of corresponding static paths in the two cases, and consequently no theoretical ground for the spectral resolution. Secondly, we shall show that the extended form of the restrictions leads to the simple Zeeman effect.

§ 2.

We consider the motion of an electron in the neighbourhood of a heavy nucleus and in the presence of a steady magnetic field H. If W is the sum of kinetic and potential energies we have* in spherical polars (r, θ, ψ):

$$E_{kin}+E_{pot} = m_0c^2\left(\frac{1}{\sqrt{(1-\beta^2)}}-1\right)-\frac{eE}{r} = W,$$

or

$$\frac{1}{\sqrt{(1-\beta^2)}} = 1+\frac{W+eE/r}{m_0c^2}, \tag{1}$$

where m and $(-e)$ are the "proper mass" and charge of the electron respectively, E the charge on the nucleus, c the velocity of light, and $c\beta$ the velocity of the electron. Also

$$\beta^2 = \frac{1}{c^2}(\dot{r}^2+r^2\dot{\theta}_2+r^2\sin^2\theta\,\dot{\psi}^2) = \frac{1}{c^2m^2}\left(p_r^2+\frac{1}{r^2}p_\theta^2+\frac{1}{r^2\sin^2\theta}p_\psi^2\right),$$

where

$$p_r = m\dot{r}, \quad p_\theta = mr^2\dot{\theta}, \quad p_\psi = mr^2\sin^2\theta\dot{\psi}, \tag{2}$$

are the generalised momenta, m being the mass (variable) of the electron, so that

$$m = m_0/\sqrt{(1-\beta^2)}. \tag{3}$$

We thus have, using (2) and (3),

$$\frac{1}{1-\beta^2} = 1+\frac{1}{c^2m^2_0}\left(p_r^2+\frac{1}{r^2}p_\theta^2+\frac{1}{r^2\sin^2\theta}p_\psi^2\right). \tag{4}$$

And from (1) and (4)

$$p_r^2+\frac{1}{r^2}p_\theta^2+\frac{1}{r^2\sin^2\theta}p_\psi^2 = 2m_0W+\frac{2m_0eE}{r}+\frac{1}{c^2}\left(W+\frac{eE}{r}\right)^2, \tag{5}$$

---

* Neglecting the ratio of m_0 to the mass of the nucleus.

which defines the total energy of the motion in terms of the Hamiltonian co-ordinates. Now*

$$\frac{d}{dt}(p_i - ea_i) + \partial W/\partial q_i = -e \sum_s \frac{\partial a_s}{\partial q_i}\frac{dq_s}{dt}, \tag{6}$$

where a is the generalised magnetic vector potential given by

$$\partial a_3/\partial q_2 - \partial a_2/\partial q_3 = g_2 g_3 \mathrm{H}_1/c, \text{ and two similar equations,} \tag{7}$$

$g\,dq$ being lineal. On applying (6) to the co-ordinate ψ we get

$$\frac{d}{dt}(p_\psi - ea_\psi) = 0, \qquad \text{so that} \qquad p_\psi - ea_\psi = \mathrm{F}, \tag{8}$$

where F is a constant. Or since

$$ea_\psi = e\mathrm{H}r^2 \sin^2\theta/2c = m_0\omega r^2 \sin^2\theta, \tag{9}$$

where

$$\omega = \tfrac{1}{2}e\mathrm{H}/m_0 c, \tag{10}$$

we have from (8) and (9)

$$p_\psi = \mathrm{F} + m_0\omega r^2 \sin^2\theta. \tag{11}$$

Substituting from (11) and (5) we have

$$r^2[-p_r^2 + 2m_0\mathrm{W} + (2m_0 e\mathrm{E})/r + 1/c^2(\mathrm{W} + e\mathrm{E}/r)^2 - 2m_0\mathrm{F}\omega]$$
$$= p_\theta^2 + \mathrm{F}^2/\sin^2\theta + m_0^2\omega^2 r^4 \sin^2\theta. \tag{12}$$

And in this last equation the variables are separable if we neglect the term in ω^2: thus

$$r^2[-p_r^2 + 2m_0\mathrm{W} + (2m_0 e\mathrm{E})/r + \frac{1}{c^2}(\mathrm{W} + e\mathrm{E}/r)^2 - 2m_0\mathrm{F}\omega]$$
$$= p^{\theta 2} + \mathrm{F}^2/\sin^2\theta = p^2 \text{ (say),†} \tag{13}$$

giving

$$p_r = \sqrt{(\mathrm{A} + 2\mathrm{B}/r + \mathrm{C}/r^2)}, \tag{14}$$

$$p_\theta = \sqrt{(p^2 - \mathrm{F}^2/\sin^2\theta)}, \tag{15}$$

where

$$\left.\begin{aligned} \mathrm{A} &= 2m_0\mathrm{W}(1 + \mathrm{W}/2m_0c^2 - \mathrm{F}\omega/\mathrm{W}) \\ \mathrm{B} &= e\mathrm{E}m_0(1 + \mathrm{W}/m_0c^2) \\ \mathrm{C} &= -(p^2 - e^2\mathrm{E}^2/c^2) \end{aligned}\right\}. \tag{16}$$

§ 3.

If now we assume for the quantum restrictions their older form, viz.,

$$\int_0 p_i\,dq_i = n_i h, \qquad i = 1, 2, 3 \ldots \tag{17}$$

---

* See G. A. Schott, *loc. cit.*, p. 292.

† As will be seen from equation (27A) below, p is the constant angular momentum in the plane of the original motion (before the field is impressed).

where the integration extends over the period of variation of q_i, we have from (11), (14), and (15):

$$\left.\begin{aligned}&\int_0 \sqrt{(A+2B/r+C/r^2)}\,dr = n_1h \qquad (\alpha)\\ &\int_0 \sqrt{(p^2-F^2/\sin^2\theta)}\,d\theta = n_2h \qquad (\beta)\\ &\int_0^{2\pi}(F+m_0\omega r^2\sin^2\theta)\,d\psi = n_3h \qquad (\gamma)\end{aligned}\right\} \tag{18}$$

where the integration in the first two cases extends from the minimum value to the maximum value, and back again to the minimum value of the two respective independent variables. The first equation in (18) yields, by the method of contour integration in the complex plane,*

$$2\pi i(\sqrt{C}+B/\sqrt{A}) = n_1h. \tag{19}$$

(18 β) easily yields by elementary methods, on writing $x = \cos\theta$,

$$p-F = n_2h/2\pi. \tag{20}$$

For the third integral we have to evaluate

$$I = \int_0^{2\pi} m_0\omega r^2\sin^2\theta\,d\psi = \omega\int_0^{\tau}\sqrt{(1-\beta^2)}\,p_\psi\,dt$$

by (2) above, where τ is the time during which ψ changes from 0 to 2π. On substituting for p_ψ from (11) and neglecting ω^2, we have

$$I = \omega F\int_0^{\tau}\sqrt{(1-\beta^2)}\,dt.$$

Let τ_0 be the value of τ calculated for $\omega = 1/c = 0$, *i.e.*, the value calculated in the absence of the field and without taking account of the relativity correction; we see that

$$I = \omega F(\tau_0 + \text{terms in } \omega \text{ and } 1/c),$$

so that, neglecting higher powers of ω than the first and products of ω and $1/c$, we have

$$I = \omega F\tau_0. \tag{21}$$

Also
$$\tau_0 = \frac{\pi a_0 b_0}{\frac{1}{2}p_0/m_0}, \tag{22}$$

$\frac{1}{2}p_0/m_0$ being the areal velocity of the electron in the elliptic path, defined in the absence of the field, whose major and minor semi-axes are a_0 and b_0 respectively; these constants could be calculated from our analysis by putting $\omega = 1/c = 0$ everywhere. The values thus obtained are identical with those obtained by Sommerfeld* by simpler methods, thus:—

---

* See Sommerfeld, *loc. cit.*, p. 477; our $\sqrt{C}$ corresponds to his $(-\sqrt{C})$.

$$\left.\begin{aligned} a_0 &= h^2(n_1+n)^2/[(2\pi)^2 m_0 e\mathrm{E}] \\ b_0 &= [h^2 n\,(n+n_1)]/[(2\pi)^2 m_0 e\mathrm{E}] \\ p_0 &= nh/2\pi \end{aligned}\right\}, \tag{23}$$

where

$$n = n_2 + n_3, \tag{24}$$

so that we have finally for the third quantum restriction

$$\mathrm{F} = n_3 h(1-\kappa)/2\pi, \tag{25}$$

where

$$\kappa = [\omega h^3 (n+n_1)^3]/[(2\pi)^3 m_0 e^2 \mathrm{E}^2]. \tag{26}$$

And from (20) and (25)

$$p = nh\,(1-\kappa n_3/n)/2\pi. \tag{27}$$

In order to calculate the energy W of the static paths, let

$$\mathrm{W} = \mathrm{W}_0 + \underset{\omega}{\delta\mathrm{W}} + \underset{c}{\delta\mathrm{W}}, \tag{28}$$

where W_0 is the value obtained for $\omega = 1/c = 0$, and $\underset{\omega}{\delta\mathrm{W}}$ and $\underset{c}{\delta\mathrm{W}}$ being first order terms in ω and $1/c^2$ respectively; and adopt a similar notation for A, B and C. We see from (16) that

$$\left.\begin{aligned} &\mathrm{A}_0 = 2m_0\mathrm{W}_0,\ \underset{\omega}{\delta\mathrm{A}} = 2m_0\left(\underset{\omega}{\delta\mathrm{W}} - \frac{n_3 h\omega}{2\pi}\right), && \underset{c}{\delta\mathrm{A}} = 2m_0\mathrm{W}_0\left(\frac{\mathrm{W}_0}{2m_0c^2} + \frac{\underset{c}{\delta\mathrm{W}}}{\mathrm{W}_0}\right) \\ &\mathrm{B}_0 = e\mathrm{E}m_0, \quad \underset{\omega}{\delta\mathrm{B}} = 0, && \underset{c}{\delta\mathrm{B}} = \frac{e\mathrm{E}\mathrm{W}_0}{c^2} \\ &\mathrm{C}_0 = -\frac{n^2h^2}{(2\pi)^2},\ \underset{\omega}{\delta\mathrm{C}} = \frac{2h^2\kappa n n_3}{(2\pi)^2}, && \underset{c}{\delta\mathrm{C}} = \frac{e^2\mathrm{E}^2}{c^2} \end{aligned}\right\} \tag{29}$$

where only first-order small terms are considered. W_0 is obtained from (19) and (27) on putting $\omega = 1/c = 0$. Thus:—

$$\mathrm{W}_0 = \frac{-\mathrm{N}h\,(\mathrm{E}/e)^2}{(n_1+n)^2}, \tag{30}$$

where N is the Rydberg constant:

$$\mathrm{N} = 2\pi^2 e^4 m_0/h^3. \tag{31}$$

The values of the two terms on the right-hand side of (19) may now be written to the first order, using (29):—

$$2\pi i\sqrt{\mathrm{C}} = -nh\left[1 - \frac{(2\pi)^2 e^2\mathrm{E}^2}{2n^2h^2c^2} - \frac{\kappa n_3}{n}\right]. \tag{32}$$

$$2\pi i\frac{\mathrm{B}}{\sqrt{\mathrm{A}}} = (n+n_1)h\left[1 + \frac{\mathrm{W}_0}{c^2m_0} - \frac{1}{2\mathrm{W}_0}\left(\underset{\omega}{\delta\mathrm{W}} - \frac{n_3h}{2\pi}\omega\right) - \tfrac{1}{2}\left(\frac{\mathrm{W}_0}{2m_0c^2} + \frac{\underset{c}{\delta\mathrm{W}}}{\mathrm{W}_0}\right)\right], \tag{33}$$

---

* *Loc. cit.*, p. 267. Sommerfeld's n' corresponds to our n_1.

so that we have from (19), on equating the terms in ω to zero from the two sides of the equation, using (32) and (33),

$$\underset{\omega}{\delta}\mathrm{W} = \frac{n_3 h\omega}{2\pi} + \frac{2\mathrm{W}_0 \kappa n_3}{n+n_1},$$

which yields, on substituting for κ and W_0 from (26) and (30) respectively,

$$\underset{\omega}{\delta}\mathrm{W} = 0, \qquad (34)$$

so that *the energy of any static path defined in the presence of the field by the quantum numbers* n_1, n_2, n_3, *would be the same as that of the corresponding path defined in the absence of the field by the same quantum numbers.* This would obviously lead to no Zeeman effect at all.

§ 4.

Instead of (17), let us now maintain the slightly extended form

$$\int_0 (p_i - ea_i)\,dq_i = n_i h, \quad i = 1, 2, 3\ldots, \qquad (17\text{A})$$

where the integration extends as before over the period of variation of q_i. Proceeding in a similar manner as in § 3, and remembering that $a_r = a_\theta = 0$ and that a_ψ is given by (9), we now have

$$\int_0 (p_\psi - m_0 \omega r^2 \sin^2\theta)\,d\psi = n_3 h, \qquad (17\text{B})$$

or

$$\int_0^{2\pi} \mathrm{F}\,d\psi = n_3 h, \qquad (18\gamma\text{A})$$

so that F and p are now given respectively by

$$\mathrm{F} = n_3 h/2\pi \qquad (25\text{A})$$

and

$$p = nh/2\pi, \qquad (27\text{A})$$

equations (19), (20) and (24) being still maintained. And adopting the same notation as before, we see that (29) still holds good except for the value of $\underset{\omega}{\delta}\mathrm{C}$, which is now given by

$$\underset{\omega}{\delta}\mathrm{C} = 0:$$

This leads, instead of (32), to the equation

$$2\pi i\sqrt{\mathrm{C}} = -nh\left[1 - \frac{(2\pi)^2 e^2 \mathrm{E}^2}{2n^2 h^2 c^2}\right], \qquad (32\text{A})$$

(30), (31) and (33) being unaffected. So that, finally, on equating the terms in ω and $1/c^2$ respectively to zero from the two sides of (19), using (32A) and (33), we have

$$\underset{\omega}{\delta}\mathrm{W} = n_3 h\omega/2\pi, \tag{34A}$$

$$\underset{c}{\delta}\mathrm{W} = \frac{-2\mathrm{W}_0{}^2}{m_0 c^2}\left[\tfrac{1}{4}+\frac{n_1}{n}\right] = \frac{-\mathrm{N}\gamma h\,(\mathrm{E}/e)^4}{(n_1+n)^4}\left[\tfrac{1}{4}+\frac{n_1}{n}\right], \tag{35}$$

where γ is a small quantity given by

$$\gamma = 4\pi^2 e^4/c^2 h^2. \tag{36}$$

§ 5.

The application of equation (34*a*) to the hydrogen atom, the singly ionised helium atom, the doubly ionised lithium atom, etc., leads in each case to a change of frequency

$$\underset{\omega}{\delta}\nu = \omega\,(m_3-n_3)/2\pi, \tag{37}$$

where m refers to a path of greater energy than the n-path. And, further, on applying Bohr's correspondence principle, we have

$$m_3-n_3 = 0 \quad \text{or} \quad \pm 1, \tag{38}$$

corresponding to the p- and n-components respectively. So that

$$\left.\begin{array}{ll} \underset{\omega}{\delta}\nu = 0 & \text{for } p\text{-components} \\ \text{or } \pm e\mathrm{H}/4\pi m_0 c & \text{for } n\text{-components} \end{array}\right\}. \tag{39}$$

The formula (35) leads to the ordinary fine structure of the spectral "line" in question, and we see that the effect of the field is to split up each component of the fine structural group into a simple Zeeman triplet.

APPENDIX.

To show that if $\Sigma_{i,1}$, $\Sigma_{i,2}$, ... $\Sigma_{i,s}$, ..., etc., and Σ_1, Σ_2, ... Σ_s, ..., etc., are the paths defined by the extended quantum restrictions, in the absence and in the presence of the field respectively, then the relation between $\Sigma_{i,s}$ and Σ_s is what it would be on classical dynamics, if an electron moving along $\Sigma_{i,s}$ were to be made to move along Σ_s by the introduction of the magnetic field.

The equations defining any static path Σ_s are obtained from the results of § 2 of the text as follows: We have from (2), (14) and (15),

$$\left.\begin{array}{l} p_r = m\dot{r} = \sqrt{(\mathrm{A}+2\mathrm{B}/r+\mathrm{C}/r^2)} \\ p_\theta = mr^2\dot{\theta} = \sqrt{\{p^2-(\mathrm{F}^2/\sin^2\theta)\}} \end{array}\right\}. \tag{i}$$

The first of these questions gives

$$\frac{dt}{dr} = \frac{m_0}{\sqrt{(1-\beta^2)}} \times \frac{1}{\sqrt{(\mathrm{A}+2\mathrm{B}/r+\mathrm{C}/r^2)}},$$

and on substituting the value of $1/\sqrt{(1-\beta^2)}$ from equation (1) of the text and integrating, we get

$$t = \int_{r_0}^{r} \frac{m_0 + \frac{W+eE/r}{c^2}}{\sqrt{(A+2B/r+Cr^2)}}\, dr \tag{ii}$$

where r_0 corresponds to $t=0$.

Further, on making the substitution

$$\cos\theta = \sin\alpha\cos\phi, \tag{iii}$$

where

$$\cos\alpha = F/p, \tag{iv}$$

we have

$$\dot\phi = \frac{\sin\theta\,\dot\theta}{\sin\alpha\,\sin\phi},$$

so that

$$mr^2\dot\phi = \frac{mr^2\sin\theta\,\dot\theta}{\sqrt{(\sin^2\theta-\cos^2\alpha)}} \text{ from (iii) } = p\times\frac{mr^2\,\dot\theta}{\sqrt{\{p^2-(F^2/\sin^2\theta)\}}} \text{ from (iv).}$$

Or, on using the second equation of (i) we see that

$$mr^2\dot\phi = p,$$

which, with the first equation of (i), gives

$$r^2\frac{d\phi}{dr} = \frac{p}{\sqrt{(A+2B/r+C/r^2)}} \quad\text{or}\quad \phi = \int_{r_0}^{r}\frac{p\,dr}{r^2\sqrt{(A+2B/r+C/r^2)}}. \tag{v}$$

We also have from (2) and (11) of the text

$$p_\psi = mr^2\sin^2\theta\dot\psi = F + m_0\omega r^2\sin^2\theta.$$

And on eliminating r^2 between this and the second equation of (i) we get

$$\dot\psi - \omega\sqrt{(1-\beta^2)} = \frac{F}{\sin^2\theta\sqrt{\{p^2-(F^2/\sin^2\theta)\}}},$$

which yields on using (iv)

$$\dot\psi - \omega\sqrt{(1-\beta^2)} = \frac{\dot\theta}{\sin^2\theta\sqrt{(\tan^2\alpha-\cot^2\theta)}}.$$

If we introduce the variable time co-ordinate t' defined by

$$dt'/dt = \sqrt{(1-\beta^2)}, \tag{vi}$$

we now have

$$\frac{d\psi}{dt'} - \omega = \frac{d\theta/dt'}{\sin^2\theta\sqrt{(\tan^2\alpha-\cot^2\theta)}};$$

which gives on integration

$$\cos(\psi-\omega t'-\Delta) = \cot\alpha\cot\theta, \tag{vii}$$

where Δ is an arbitrary constant depending on the values of ψ and θ at $t=0$. The five equations (iii) to (vii) define the path of the electron, whilst equation (ii) defines the mode in which it is described for any given set of values of the five constants involved, viz., W, F, p, r_0, and Δ. The equations defining the path $\Sigma_{i,s}$, in the absence of the field, are clearly obtained from the above equations defining Σ_s on putting $\mathrm{H} = 0$ everywhere, and writing for the constants involved a corresponding set of values (denoted by the suffix i). We thus have for $\Sigma_{i,s}$

$$\left.\begin{aligned} t = \int_{r_{0,i}}^{r} \frac{m_0 + \mathrm{W}_i + e\mathrm{E}/r}{\sqrt{(\mathrm{A}_i + 2\mathrm{B}_i/r + \mathrm{C}_i/r^2)}}\,dr, \quad \phi = \int_{r_{0,i}}^{r} \frac{p_i\,dr}{r^2\sqrt{(\mathrm{A}_i + 2\mathrm{B}_i/r + \mathrm{C}_i/r^2)}} \\ \cos(\psi - \Delta_i) = \cot\alpha_i \cot\theta \\ \text{where} \qquad \cos\theta = \sin\alpha_i\cos\phi, \qquad \cos\alpha_i = \mathrm{F}_i/p_i. \end{aligned}\right\} \quad \text{(viii)}$$

Now the extended quantum restrictions define the relations between some of these constants according to equations (25A), (27A), and (34A), thus

$$p = p_i = nh/2\pi; \quad \mathrm{F} = \mathrm{F}_i = n_3h/2\pi; \quad \mathrm{W} = \mathrm{W}_i + \omega\mathrm{F}_i. \qquad \text{(ix)}$$

So that, remembering the meanings of A, B, C given by equations (16) of the text and neglecting (as in the text) terms in ω/c^2, and in ω^2 and higher powers, we have on applying (ix)

$$\mathrm{A} = \mathrm{A}_i, \quad \mathrm{B} = \mathrm{B}_i, \quad \mathrm{C} = \mathrm{C}_i, \quad \alpha = \alpha_i, \quad \frac{\mathrm{W} + e\mathrm{E}/r}{c^2} = \frac{\mathrm{W}_i + e\mathrm{E}/r}{c^2}. \qquad \text{(x)}$$

We may further choose our origins of time, in the presence and absence of the field respectively, such that

$$r_0 = r_{0,i}. \qquad \text{(xi)}$$

So that we finally see from (vii), (viii), (x) and (xi) that the set of equations (viii) which define the complete motion along $\Sigma_{i,s}$, would also define the motion along Σ_s if we wrote χ for ψ, where

$$\chi = \psi - \omega t' + (\Delta_i - \Delta). \qquad \text{(xii)}$$

In other words, the motion along Σ_s is obtained from the motion along $\Sigma_{i,s}$ by the mere superposition of an angular velocity about Oz equal to

$$\dot{\psi} - \dot{\chi} = \omega\,dt'/dt = \omega\sqrt{(1-\beta^2)}. \qquad \text{(xiii)}$$

This, of course, is merely a statement of Schott's theorem referred to in the text. We have here established it *as a necessary result of the extended quantum restrictions*, and without considering the period of establishment of the field.

The Magnitude of the Gyromagnetic Ratio.

By O. W. Richardson, F.R.S., King's College, London.

(Received December 4, 1922.)

The accurate experiments of Chattock and Bates[*] prove that the angular momentum arising in a ferromagnetic substance from unit change in its magnetic moment is very nearly, if not exactly, one half the value $2m/e = 1 \cdot 13 \times 10^{-7}$, which seemed to me the most likely when I first discussed this effect.[†] This conclusion is supported by the fact that the improvements which have been introduced into this subject by successive experimenters[‡] in recent years have led to values showing a strong tendency to settle at the same limit $m/e = 5 \cdot 65 \times 10^{-8}$. This value is also in general agreement with that deduced by Barnett[§] from experiments on the converse effect. It seems desirable therefore to reconsider the interpretation of this ratio.

The higher value $2m/e$ is obtained by making rather definite assumptions, which evidently require modification, as to the nature of the phenomena. These assumptions are that the process of magnetization involves the turning of electron orbits, and that nothing else which may occur has any important influence on the phenomena. The inertia of the electrons is assumed to be entirely of the type which controls the deflection of a beam of cathode rays by a magnetic field, and any change in the motion of the positively charged part of the atom is disregarded. These assumptions are essentially the same as those of the theories of Langevin and Weiss which have been successful in dealing with purely magnetic phenomena.

One way of escape is to suppose that the inertia of the electrons in these magnetic phenomena is not of the usual type. As an illustration of such possibilities, I showed at the Solvay Conference at Brussels in 1921 that the spin imposed on an electron by placing it in a magnetic field gives rise to an angular momentum and magnetic moment whose ratio depends on the constitution assumed for the electron. If the electron is assumed to be a charged sphere, the charges being held in position by a non-electromagnetic tension of the amount suggested by Poincaré, the ratio of angular momentum to magnetic moment created is found to be m/e instead of $2m/e$. This effect, however, is diamagnetic and very small, and, in fact, such ways of escape do not seem at all promising. The turning-orbit theory of the Langevin type goes so far in explaining magnetic phenomena that it is almost certain that it is true in its essentials.

---

[*] Chattock and Bates, 'Phil. Trans.,' A, vol. 223, p. 257 (1922).

[†] Richardson, 'Phys. Rev.' (1), vol. 26, p. 248 (1908).

[‡] Einstein and de Haas, 'Verh. d. Deutsch. Physik. Ges.,' vol. 17, p. 152 (1915). De Haas, 'Verh. d. Deutsch. Physik. Ges.,' vol. 18, p. 423 (1916). J. Q. Stewart, 'Phys. Rev.' (2), vol. 11, p. 100 (1918). Beck, 'Ann. der Physik,' (4), vol. 60, p. 109 (1919). Arvidsson, 'Physik. Zeits.,' vol. 21, p. 88 (1920).

[§] Barnett, 'Phys. Rev.' (2), vol. 6, p. 239 (1915); vol. 10, p. 7 (1917).

We are now practically driven, I think, to the conclusion that the motions of the positively charged parts of the atoms cannot be disregarded in considering these gyromagnetic phenomena. The formulæ relevant to their discussion are given in my paper in the 'Physical Review' in 1908. They are

$$U_z = 2\,(\mathrm{NM}\bar{\mathrm{A}} + nm\bar{a}) \qquad (1)$$

and

$$M_z = \mathrm{NE}\bar{\mathrm{A}} + ne\bar{a}, \qquad (2)$$

where U_z and M_z are respectively angular momentum and magnetic moment per unit volume for any axis Oz, N, n numbers of electrons per unit volume, M, m masses of electrons and $\bar{\mathrm{A}}$, $\bar{a}$ the respective mean areal velocities about the Oz axis: capital letters refer to positive and small letters to negative electrons. If we assume the effect of the areal velocity of the positive parts to be negligible and put $\bar{\mathrm{A}} = o$ in these formulæ we clearly get the ratio of angular momentum to magnetic moment to be $2m/e$, the value appropriate to turning orbits of negative electrons. If the areal velocity of the positive parts is not negligible it is clear that it will be much more important in the expression (1) for the angular momentum than it is in expression (2) for the magnetic moment. This follows because the numerical magnitude of NM/nm is much greater than that of NE/ne. In the unequivocal case of the hydrogen atom NM/nm is about 1800 times NE/ne. With the ferromagnetic elements the values of these quantities depend on the precise allocation we give to the magnetically effective orbits, but a disparity of comparable magnitude will always be found to subsist. To obtain a value of the ratio U_z/M_z which is smaller than $2m/e$ it is necessary that the averaged contribution of the positive charges to the angular momentum established by magnetization should be oppositely directed to that of the negative charges.

This does not necessarily mean that the positive particles are revolving in the atom in an opposite sense to those of the negative electrons, both sets might be revolving in the same sense, but in the act of magnetization the normals to one set might tend to align with the direction of the applied field and the other to move away from it. They would, in fact, so tend, owing to the opposite magnetic polarity of orbits executed in the same sense by charged particles of opposite sign. For the ratio U_z/M_z to fall to the experimental value m/e it is necessary and sufficient, on this view of the phenomena, that in the establishment by magnetization of angular momentum the contribution arising from the positively charged parts of the atom should always be one half of that from the negative electrons and oppositely directed to it. This statement is exact to the degree of accuracy which disregards the effect on the magnetic moment of the positively charged parts in comparison with that from the negatively charged parts of the atom.

It seems possible to proceed further with the help of Bohr's theory. According to Ewing* the saturation intensity of magnetization of iron is very nearly 2×10^{-20} c.g.s. units per atom of iron. The angular momentum corresponding to this agrees within a few per cent. with twice the fundamental quantum unit $h/2\pi$. This strongly suggests that the magnetism of iron is due to a pair of parallel or coplanar orbits each having a single azimuthal quantum. I am not able to find that Bohr has published a detailed picture of the structure of the iron atom, but from the constitutions proposed for argon and krypton it is clear that it will have a number of pairs of symmetrical orbits each with one azimuthal quantum. To account for the fact that the ratio U_z/M_z is m/e instead of $2m/e$ it is only necessary to assign another azimuthal quantum to the nucleus. If this orbit turns with the field in the same sense as the two effective electron orbits, as the quantum conditions seem to require, its angular momentum will have to be oppositely directed to theirs. In this case, since $nma = -2\text{MNA}$, the A's now denoting the actual areal velocities to which the mean values have become proportional, we have $\text{NA}/na = -m/2\text{M}$ and

$$U_z/M_z = \frac{m}{e}\Big/\left(1-\frac{1}{2}\frac{E}{e}\frac{m}{M}\right). \tag{3}$$

For the iron nucleus (atomic number 26) $E = -26e$ and $M = 56 \times 1835m$. If these numbers are substituted in (3) it is found that the ratio is less than m/e by a little more than one part in 10,000. The denominator on the right-hand side of (3) represents the effect of the motion of the nucleus on the magnetic moment, and the fact that it differs so slightly from unity justifies our neglect of this effect in discussing the broad features of the phenomena. There are a number of other possible arrangements on the quantum theory which would fit the facts, but this seems to be the simplest and most probable.

---

* 'Proc. R. S. Edin.,' vol. 42, p. 124 (1922).

Stability of a Viscous Liquid contained between Two Rotating Cylinders.

By G. I. Taylor, F.R.S.

(Received March 21, 1922.)

(Abstract.)

Part I.—The stability for symmetrical disturbances of a viscous fluid in steady motion between concentric rotating cylinders is investigated mathematically. It is shown that at slow speeds the motion is always stable, but that at high speeds the motion is only stable when the ratio of the speed of the outer cylinder to that of the inner one exceeds a certain value. When the ratio is less than this or when it is negative the motion becomes unstable at high speeds. The "criterion" for stability is found, and in cases suitable for experimental verification an approximate form for the "criterion" is developed which is useful for numerical computation.

The type of instability which may be expected to appear when the speed of the cylinders is slowly increased is shown to consist of symmetrical ring-shaped vortices spaced at regular intervals along the length of the cylinders. These vortices rotate alternately in opposite directions. Their dimensions are calculated and it is shown that they are contained in partitions of rectangular cross-section. In the case when the instability arises while both cylinders are rotating in the same direction, these rectangles are squares, so that the vortices are spaced at distances apart equal to the thickness of the annular space between the two cylinders. In the case when the cylinders rotate in opposite directions the spacing, or distance between the centres of neighbouring vortices, is smaller than this; and at the same time two systems of vortices develop—an inner system which is similar to the system which appears when the two cylinders rotate in the same direction, and an outer system, which is much less vigorous and rotates in the opposite direction to the adjacent members of the inner system.

Part II.—In Part II experiments are described in which all the conclusions of Part I are subjected to numerical verification. It is shown that the predicted vortices do in fact appear at the predicted speeds, and photographs of the vortices are shown. The accuracy with which the critical speeds can be predicted is good, the error being about 2 per cent. This compares favourably with the best previous criterion for stability, namely, that of Reynolds,

modified by Orr, which, if it may be regarded as a "criterion" at all, is in error by 7000 per cent.

The error in the predicted spacing of the vortices is 3 to 5 per cent., and the rectangular partitions which contain them are made visible in the photographs.

At the end of the paper some photographs are shown and observations described which throw some light on the way in which the initial symmetrical instability begins to break down before the liquid becomes completely turbulent. The results are contrary to those of previous observers, partly because special methods were designed, with a view to showing up the particular type of motion which theory leads one to expect, but more specially because of the care which was taken to eliminate the effects of the ends of the cylinders.

Discontinuous Fluid Motion Past Circular and Elliptic Cylinders.

By S. Brodetsky, M.A., Ph.D., University, Leeds.

(Communicated by Prof. G. H. Bryan, F.R.S.—Received June 29, 1922.)

1. The problem of discontinuous fluid motion past a curved barrier has become one of the classical problems of mechanics. For plane barriers the Schwartz-Christoffel transformation offers a direct method of solution; in the case of curved barriers no direct method of solution has been found. "A quicker start . . . can be made . . . by the simpler process of writing down a likely expression . . . and then investigating the streaming motion implied and the shape of the boundary."*

In the present paper the procedure is essentially in pursuance of Greenhill's advice, with the important modification that an attempt is made to solve problems that have a bearing on practical applications. Although many types of barriers have been suggested by various writers, very little in the way of actual numerical calculation has been done. Apparently, the only case worked out in detail is that by Brillouin,† and this case is somewhat artificial and of little use. Such barriers as the circular and elliptic cylinders—important for the aerodynamics of struts—have not been attempted at all, while application to a barrier like a modern aeroplane wing seems very remote indeed.

---

* Sir G. Greenhill, Appendix to Report 19 (1910), Advisory Committee for Aeronautics (1916).

† 'Ann. de Chim. et Phys.,' vol. 23, pp. 145-230 (1911).

It is true that exact mathematical solutions of such problems seem at present to be quite unobtainable. Approximate solutions can, however, be obtained to any required degree of accuracy, and our object here is to show how the method is applied, and also to get results within a 1-per-cent. degree of accuracy in the case of circular and elliptic cylinders. The method is not really new, since it is based on the transformation introduced by Levi-Civita,* and described in some detail in Brillouin's paper; a valuable account of this transformation, with extensions, is also to be found in U. Cisotti's 'Idromeccanica Piana,' Milano, 1921-2.

We consider here the symmetrical problem, *i.e.*, the case in which the barrier is in the form of a cylinder with a plane of symmetry, and the direction of relative motion through the fluid is parallel to the plane of symmetry and perpendicular to the generators of the barrier. We make the problem "real" and of practical use by assuming that the free stream-lines leave the barrier with finite curvature. Since the problem is symmetrical, we know where the fluid divides into the two parts going round the two sides of the barrier; we do not know, however, in the case of a real barrier, where the free stream lines leave the barrier—we can only say by common sense that they must do so before the barrier becomes parallel to the plane of symmetry.

2. The motion is two-dimensional, and our notation is as follows: The complex variable $z\,(\equiv x+\iota y)$ gives the geometry of the actual motion in any plane perpendicular to the generators of the barrier. The velocity components are defined in terms of the potential and stream functions:—

$$u = \partial\phi/\partial x = \partial\psi/\partial y, \quad v = \partial\phi/\partial y = -\partial\psi/\partial x.$$

We use the complex variable $w\,(\equiv \phi+\iota\psi)$ and define

$$\zeta \equiv re^{\iota\theta} = dz/dw\,; \quad \Omega = \log\zeta = \log r + \iota\theta.$$

The problem is to find w in terms of z, the method usually adopted being to find ζ or Ω in terms of w. Fig. 1 shows the z, w, ζ and Ω planes with the boundary conditions, the barrier being taken at rest and the fluid streaming past it with given velocity at infinity. IC is in the plane of symmetry: CA, CA′ are the stream-lines in contact with the barrier; AJ, A′J′ are the free stream-lines. We choose the velocity at infinity to be unity, and also take CA in the w plane to be of unit length. We now introduce Levi-Civita's transformation in a modified form, namely,

$$\sqrt{w} = (\tau - 1/\tau)/2\iota. \tag{1}$$

---

* 'Rend. del Circ. Mat. di Palermo,' vol. 23 (November, 1906).

It is easily verified that in the τ plane the part of the barrier between A, A′ is represented by the semicircle ACA′, while the free stream-lines correspond to the two halves of the diameter AA′ along the imaginary axis.

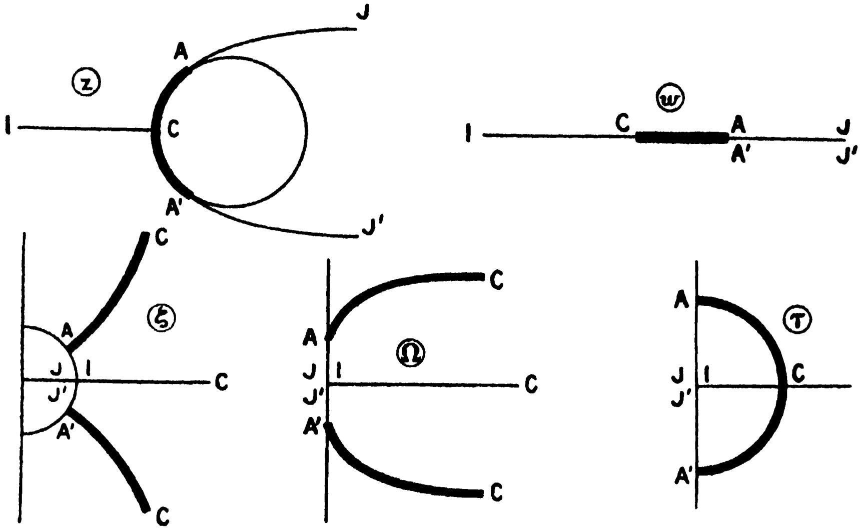


FIG. 1.

When the barrier is plane the relation between Ω, τ is

$$\Omega = \log(1+\tau)/(1-\tau).$$

In the Bobyleff problem, where the barrier consists of two planes, CA, CA′, at an angle $\lambda\pi$, the form of Ω is

$$\Omega = \lambda \log(1+\tau)/(1-\tau).$$

For the general symmetrical problem we use

$$\Omega = \lambda \log(1+\tau)/(1-\tau)+\Omega',$$

where Ω' is some function of τ, finite for $|\tau| \leq 1$. Since Ω must be purely imaginary for $\tau = \iota\rho$, $-1 \leq \rho \leq 1$, and also $\Omega(-\iota\rho) = -\Omega(\iota\rho)$, we conclude that Ω' must be expressible in the form of a Taylor expansion in odd powers of τ. Hence

$$\Omega = \lambda \log(1+\tau)/(1-\tau)+A_1\tau+\tfrac{1}{3}A_3\tau^3+\tfrac{1}{5}A_5\tau^5+\ \ldots. \qquad (2)$$

We use the notation $\tau = \rho e^{\iota\sigma}$ and points on ACA′ are given by $\rho = 1$, $-\pi/2 \leq \sigma \leq \pi/2$. Points on the free stream-lines are given by $\tau = \iota\rho$, $-1 \leq \rho \leq 1$, as already used.

For points on CA we get from (2)

$$\theta = \tfrac{1}{2}\lambda\pi + A_1 \sin\sigma + \tfrac{1}{3}A_3 \sin 3\sigma + \tfrac{1}{5}A_5 \sin 5\sigma + \ldots.$$

Thus, as $\sigma \to 0$, $\theta \to \lambda\pi/2$, so that λ is defined as the angle at the nose C, divided by π, this angle being measured on the side of C away from the moving fluid. If the barrier is blunt at C we have $\lambda = 1$, and we shall deal only with such barriers here. Problems with $\lambda \neq 1$ can be treated similarly. We thus have

$$\Omega = \log \zeta = \log r + \iota\theta = \log (1+\tau)/(1-\tau) + A_1\tau + \tfrac{1}{3}A_3\tau^3 + \tfrac{1}{5}A_5\tau^5 + \ldots. \quad (3)$$

3. Consider the curvature of the free stream-line at A. For $\tau = \iota\rho$ we get

$$r = 1, \quad \theta = 2\tan^{-1}\rho + A_1\rho - \tfrac{1}{3}A_3\rho^3 + \tfrac{1}{5}A_5\rho^5 - \ldots.$$

But if ds is an element of length in the z plane, we have

$$ds/d\phi = r.$$

Also, by (1),

$$\phi = \tfrac{1}{4}(\rho + 1/\rho)^2; \quad \text{hence } ds/d\rho = -(1-\rho^4)/2\rho^3.$$

Further,

$$d\theta/d\rho = 2/(1+\rho^2) + A_1 - A_3\rho^2 + A_5\rho^4 - \ldots.$$

Thus the radius of curvature of the free stream-line is given by

$$-\frac{ds}{d\theta} = \frac{(1-\rho^4)/2\rho^3}{2/(1+\rho^2) + A_1 - A_3\rho^2 + A_5\rho^4 - \ldots}.$$

The radius of curvature is therefore generally zero at A, where $\rho^2 = 1$, so the barrier ends abruptly at A, and, of course, also at A′. This cannot be the case in real problems, so that for such problems we get the condition

$$1 + A_1 - A_3 + A_5 - \ldots = 0. \quad (4)$$

If this condition is satisfied, the actual radius of curvature of the free stream-line at A is

$$2/(A_1 - 3A_3 + 5A_5 - \ldots). \quad (5)$$

This is the same as the radius of curvature of the barrier at A. For if we use $\tau = e^{\iota\sigma}$ we get for positive σ:—

$$r = \sqrt{(1+\cos\sigma)/(1-\cos\sigma)}\,.\,e^{A_1\cos\sigma + \frac{1}{3}A_3\cos 3\sigma + \ldots};$$

$$\theta = \tfrac{1}{2}\pi + A_1\sin\sigma + \tfrac{1}{3}A_3\sin 3\sigma + \ldots \quad .$$

Now $\quad ds/d\phi = r$, and $\phi = \sin^2\sigma$ by (1).

We therefore have

$$\frac{ds}{d\theta} = \frac{2r\sin\sigma}{A_1 + A_3(\cos 3\sigma/\cos\sigma) + A_5(\cos 5\sigma/\cos\sigma) + \ldots}.$$

When $\sigma \to \pi/2$, $r \to 1$, and $\cos(2n+1)\sigma/\cos\sigma \to -(2n+1)$, and so at A we get the value (5).

4. The resultant pressure on the barrier is, of course, in the plane of symmetry, and its value per unit length with unit density of the fluid is

$$\pi(1 + \tfrac{1}{2}A_1)^2 \quad (6)$$

(Brillouin, p. 195). This is the value for unit density, unit velocity at infinity, and unit length of CA in the w plane. For other values of these quantities we multiply by the first power of the density, the second power of the velocity at infinity, and the first power of the length CA in the w plane. The last defines the size as distinct from the shape.

Brillouin takes the case where all the quantities A are zero, except A_1. In order to get a real barrier the condition (4) gives $A_1 = -1$, and this is the only useful case calculated by him. We shall first show how by a judicious choice of constants A we can obtain the result for a circular cylinder to any required degree of approximation. It will save repetition if we at once say that, in order to get condition (4) automatically satisfied, we take

$$A_1 = -1+a_1,\ A_3 = a_1+a_3,\ A_5 = a_3+a_5,\ \ldots. \tag{7}$$

5. Let all the quantities a be zero, except a_1. Then

$$\zeta = (1+\tau)/(1-\tau)\,.\,e^{-(1-a_1)\tau+\frac{1}{3}a_1\tau^3};$$

so that

$$r = \sqrt{(1+\cos\sigma)/(1-\cos\sigma)}\,.\,e^{-(1-a_1)\cos\sigma+\frac{1}{3}a_1\cos 3\sigma};$$

$$\theta = \tfrac{1}{2}\pi-(1-a_1)\sin\sigma+\tfrac{1}{3}a_1\sin 3\sigma;$$

for positive values of σ.

The radii of curvature at C, A respectively are, numerically,

$$4e^{-1+\frac{4}{3}a_1}/(1-2a_1),\ 2/(1+2a_1).$$

Now choose a_1 so that these two radii are equal: we find

$$a_1 = 0{\cdot}0574,$$

and so we may expect that

$$\zeta = (1+\tau)/(1-\tau)\,.\,e^{-0{\cdot}9426\tau+0{\cdot}0191\tau^3}$$

will give an approximation to a circular cylinder. This is in fact the case, for

$$\frac{ds}{d\theta} = -\frac{2(1+\cos\sigma)}{1-0{\cdot}1147\cos 2\sigma}\,.\,e^{-0{\cdot}9426\cos\sigma+0{\cdot}0191\cos 3\sigma}$$

and we find that for values of σ between 0 and $\pi/2$ the radius of curvature is, numerically, not very different from the value at 0 or $\pi/2$, namely, 1·794. The greatest deviation occurs at about $\sigma = \pi/4$, where the value is 1·73. This is about $3\frac{1}{2}$ per cent., and we have a first approximation to a circular cylinder.

By (6) we get that the pressure per unit length with unit density and velocity is

$$R = \pi(0{\cdot}5287)^2$$

for a cylinder of diameter 3·588. If we call R_0 the pressure on a plane barrier of the same width, placed symmetrically with respect to the current, we find that approximately

$$R/R_0 = 0{\cdot}56.$$

We also find that the points A, A′, where the free stream-lines begin, are given by an arc CA corresponding to an angle

$$\tfrac{1}{2}\pi - \theta_A = 55{\cdot}1^\circ.$$

A second approximation can now be obtained. Let a_1, a_3 both exist, and choose them so that the radii of curvature are equal at $\sigma = 0$, $\sigma = \pi/4$, $\sigma = \pi/2$. The approximation already obtained reduces the laboriousness of the arithmetic, which we need not reproduce here. We get

$$a_1 = 0{\cdot}0585,\ a_3 = -0{\cdot}0083,$$

so that

$$\frac{ds}{d\theta} = -\frac{2(1+\cos\sigma)\,.\,e^{-0{\cdot}9415\cos\sigma + 0{\cdot}0167\cos 3\sigma - 0{\cdot}0017\cos 5\sigma}}{1 - 0{\cdot}1170\cos 2\sigma + 0{\cdot}0166\cos 4\sigma}.$$

The numerical value of the radius of curvature is now found to vary between the limits 1·75, 1·77, the average being 1·761. The maximum deviation is only a fraction of 1 per cent., and the barrier is now very close indeed to a circular cylinder. We get the results

$$R/R_0 = 0{\cdot}568,\ \tfrac{1}{2}\pi - \theta_A = 55^\circ.$$

Higher approximations can be sought for if desired, but in reality the labour involved would be wasted as far as practical use is concerned. Perhaps in this way the exact mathematical solution may suggest itself; the present writer has not been successful in guessing it.

6. Return to the case where only a_1 exists, and try some other value. Let us take, *e.g.*, the case $a_1 = 0$, considered by Brillouin, so that

$$\zeta = (1+\tau/1-\tau)\,.\,e^{-\tau}; \quad \text{and} \quad \theta = \tfrac{1}{2}\pi - \sin\sigma,\ ds/d\theta = -2(1+\cos\sigma)\,e^{-\cos\sigma}$$

for positive σ. The barrier is now approximately elliptic. To find an ellipse to which the barrier approximates, we take the ellipse whose vertex is at C, with an axis in the plane of symmetry, and whose radius of curvature at C is equal to that of the barrier at C; also so that where the ellipse is parallel to the barrier at A the radius of curvature is the same as that of the barrier at A. Given the radii of curvature at C and A, numerically, and the value of $\tfrac{1}{2}\pi - \theta_A$ at A, it is an easy problem to solve for the ellipse, and we need not give the details of the calculation. If a is the semi-axis lying in the plane of symmetry, and b the semi-axis perpendicular to the plane of symmetry, we find

$$a = 1{\cdot}993, \qquad b = 1{\cdot}713.$$

Calculating θ for different values of σ, and comparing the corresponding radii of curvature for the barrier and the ellipse, the results are:—

σ.	$-ds/d\theta$ for barrier.	$-ds/d\theta$ for ellipse.
0°	1·473*	1·473
30	1·570	1·615
45	1·683	1·753
60	1·819	1·881
90	2	2

The maximum difference is thus at about $\sigma = \pi/4$, being then 4 per cent. The error in the barrier is actually less. Brillouin calculates the co-ordinates of A referred to C, and finds them to be $x = 0{\cdot}777$, $y = 1{\cdot}343$. The corresponding point on the approximate ellipse has co-ordinates $x = 0{\cdot}800$, $y = 1{\cdot}373$, and the error is therefore less than 3 per cent.

The pressure is $\pi/4$ by (6), and we deduce that for an elliptic cylinder of eccentricity 0·5106 with its major axis parallel to the current, the ratio of the pressure R to the pressure R_0 on a plate of equal width, namely, 3·426, placed normally to the current, is approximately

$$R/R_0 = 0{\cdot}52.$$

Also the point where the free stream-lines leave the cylinder are given by

$$\tfrac{1}{2}\pi - \theta_A = 1^c = 57{\cdot}3^\circ.$$

We can now attempt a second approximation. The procedure is not exactly the same as for the circular cylinder, since there all that was needed was to get the radius of curvature sufficiently close to some constant value. Here we have to consider the radius of curvature as a function of θ. Our method will therefore be to keep the value of a_1 chosen, in this case 0, and choose a value of a_3, so that we get a closer approximation to an ellipse, not the same as the one already found, but one close to it. The arithmetic is easy since we can assume a_3 small.

Let us then write

$$\zeta = (1+\tau)/(1-\tau)\,.\,e^{-\tau+\frac{1}{3}a_3\tau^3+\frac{1}{5}a_3\tau^5};$$

$$\theta = \pi/2 - \sin\sigma + \tfrac{1}{3}a_3 \sin 3\sigma + \tfrac{1}{5}a_3 \sin 5\sigma;$$

$$\frac{ds}{d\theta} = -\frac{2(1+\cos\sigma)\,.\,e^{-\cos\sigma+\frac{1}{3}a_3\cos 3\sigma+\frac{1}{5}a_3\cos 5\sigma}}{1-2a_3\cos 4\sigma}.$$

If a_3 is small we can use

$$ds/d\theta = -2(1+\cos\sigma)\,e^{-\cos\sigma}(1+\tfrac{1}{3}a_3\cos 3\sigma + 2a_3\cos 4\sigma + \tfrac{1}{5}a_3\cos 5\sigma).$$

---

* The value 1·571 given by Brillouin, p. 227, is obviously wrong, due to a confusion between $4/e$ and $\pi/2$.

The presence of a_3 thus introduces correction factors in the radii of curvature; they are

$$(1+38/15\,.\,a_3) \text{ at } \sigma=0\,; \qquad (1-[30+4\sqrt{2}]/15\,.\,a_3) \text{ at } \sigma=\tfrac{1}{4}\pi\,;$$
$$(1+2a_3) \text{ at } \sigma=\tfrac{1}{2}\pi.$$

Take the mean of the first and last—since a_3 is small, we can call it the arithmetic or geometric mean, as we please—and divide all three by it. We get the values

$$1+4/15\,.\,a_3\,; \qquad 1-(64+4\sqrt{2})/15\,.\,a_3\,; \qquad 1-4/15\,.\,a_3\,;$$

and these at once suggest that if we make the second one just obtained equal to 1·04, we ought to have a second approximation. The value of a_3 is

$$a_3 = -0{\cdot}0086.$$

Since the changes introduced in θ are very small, it is hardly necessary to recalculate the approximate ellipse, and we can find it with sufficient accuracy by multiplying the dimensions of the one already found by $(1+34a_3/15)$, *i.e.*, 0·9805. If we do recalculate *ab initio* we find the ellipse

$$a=1{\cdot}958, \qquad b=1{\cdot}680\,;$$

and the comparison between the barrier and ellipse is now thus:—

σ.	$-ds/d\theta$ for barrier.	$-ds/d\theta$ for ellipse.
0°	1·441	1·441
30	1·586	1·581
45	1·717	1·720
60	1·838	1·848
90	1·966	1·966

The approximation is very good, and the results are for such an ellipse

$$R/R_0 = 0{\cdot}531, \quad \tfrac{1}{2}\pi-\theta_A = 56^\circ\,38'.$$

7. For other elliptical cylinders we start off with other values of a_1. If a_1 is between 0 and 0·0574, we get first approximations to ellipses whose major axes are in the plane of symmetry and eccentricities are between 0·5142 and 0. Negative values of a_1 are useful for such ellipses with smaller eccentricities. Values of a_1 greater than 0·0574 yield ellipses with major axes perpendicular to the current. In order to have a uniform notation we shall always use a for the semi-axis in the plane of symmetry, and b for the semi-axis perpendicular to this plane. Thus for $b/a<1$ we use a_1 algebraically less than 0·0574; for $b/a>1$ we use a_1 greater than 0·0574.

When we start off with a value of a_1 considerably more than 0·0574, or considerably less than 0, the second approximation is not sufficiently accurate. It is either necessary to go to a higher approximation, or to use some other device. We consider first an example where a higher approximation is necessary. This occurs when a_1 is considerably more than 0·0574, and as a rather extreme instance we shall start off with

$$a_1 = 0{\cdot}25.$$

We now have

$$\frac{ds}{d\theta} = -\frac{2(1+\cos\sigma).e^{-\frac{3}{4}\cos\sigma + 1/12\,.\,\cos 3\sigma}}{1-\frac{1}{2}\cos 2\sigma}; \quad \theta = \tfrac{1}{2}\pi - \tfrac{3}{4}\sin\sigma + \tfrac{1}{12}\sin 3\sigma;$$

and this is a first approximation for the ellipse

$$a = 1{\cdot}353, \qquad b = 2{\cdot}357,$$

giving $\qquad R/R_0 = 0{\cdot}59, \ \tfrac{1}{2}\pi - \theta_A = 47{\cdot}75°.$

The approximation is, however, very rough, since about $\sigma = \pi/4$ we get a maximum difference in the radius of curvature of about 22 per cent.

For the second approximation we write

$$\zeta = (1+\tau)/(1-\tau)\,.\,e^{-\frac{3}{4}\tau + \frac{1}{3}(\frac{1}{4}+a_3)\tau^3 + \frac{1}{5}a_3\tau^5},$$

and choose a_3 so that the radius of curvature at $\sigma = \pi/4$ is equal to that at the corresponding point of that ellipse which is defined by the radius of curvature at C and the radius of curvature and direction at A. This needs a certain amount of calculation and interpolation; we find

$$a_3 = -0{\cdot}0468,$$

the approximate ellipse being

$$a = 1{\cdot}241, \qquad b = 2{\cdot}047$$

giving $\qquad R/R_0 = 0{\cdot}682, \ \tfrac{1}{2}\pi - \theta_A = 47{\cdot}4°.$

The comparison of the barrier with the ellipse is now as follows :—

σ.	$-ds/d\theta$ for barrier.	$-ds/d\theta$ for ellipse.
0°	3·375	3·375
30	2·794	2·692
45	2·128	2·125
60	1·594	1·642
90	1·255	1·255

The approximation is comparatively close, and a third approximation will give a sufficiently accurate result.

To find the third approximation we write

$$\zeta = (1+\tau)/(1-\tau)\,.\,e^{-\frac{3}{4}\tau + 0{\cdot}0677\tau^3 + \frac{1}{5}(-0{\cdot}0468+a_5)\tau^5 + \frac{1}{7}a_5\tau^7},$$

so that

$$\frac{ds}{d\theta} = -\frac{2(1+\cos\sigma).e^{-\frac{3}{4}\cos\sigma + 0{\cdot}0677\cos 3\sigma + \frac{1}{5}(-0{\cdot}0468+a_5)\cos 5\sigma + \frac{1}{7}a_5\cos 7\sigma}}{1-\frac{1}{2}\cos 2\sigma + 0{\cdot}0936\cos 4\sigma - 2a_5\cos 6\sigma}.$$

The correction factors introduced by a_5, which is assumed small, are

$$(1+3{\cdot}71a_5) \text{ at } \sigma = 0°; \quad (1-3{\cdot}15a_5) \text{ at } \sigma = 30°; \quad (1-0{\cdot}0404a_5) \text{ at } \sigma = 45°;$$

$$(1+1{\cdot}83a_5) \text{ at } \sigma = 60°; \quad (1-1{\cdot}25a_5) \text{ at } \sigma = 90°.$$

By a process of trial and interpolation we find that a good result is given by

$$a_5 = 0{\cdot}0085,$$

the approximate ellipse being

$$a = 1{\cdot}233, \qquad b = 2{\cdot}072,$$

which gives

$$\mathrm{R}/\mathrm{R}_0 = 0{\cdot}673, \quad \tfrac{1}{2}\pi - \theta_{\mathrm{A}} = 47°\ 22'.$$

The difference between the radii of curvature of the barrier and ellipse is never more than 1½ per cent., and it is very probable that actually the result is correct within 1 per cent.

8. For values of b/a less than 0·858, a good approximation can be found easily by means of the following device. If we expand

$$\mu \log (1+a\tau)/(1-a\tau), \tag{8}$$

in which a is a positive number less than unity, and μ is any number, we get

$$2\mu\,(a\tau + \tfrac{1}{3}a^3\tau^3 + \tfrac{1}{5}a^5\tau^5 + \ldots).$$

It is therefore allowable to use (8) for Ω' in (2), so that

$$\zeta = \left(\frac{1+\tau}{1-\tau}\right)\left(\frac{1+a\tau}{1-a\tau}\right)^{\mu}.$$

To get a real problem condition (4) gives

$$\mu = -1/\sin\alpha, \quad \text{where} \quad \alpha = 2\tan^{-1} a,$$

i.e., α is some angle between 0 and $\pi/2$. It is clear that no generality is lost by assuming α positive. It is a fact that the form now obtained, viz.,

$$\zeta = \left(\frac{1+\tau}{1-\tau}\right)\left(\frac{1+a\tau}{1-a\tau}\right)^{-1/\sin\alpha}, \tag{9}$$

gives good first approximations for ellipses in which b/a is less than 0·858. For this value of ζ we have

$$\theta = \tfrac{1}{2}\pi - \frac{1}{\sin\alpha}\tan^{-1}(\tan\alpha\sin\sigma);$$

$$\frac{ds}{d\theta} = -\frac{2\,(1+\cos\sigma)}{\cos\alpha}\left(\frac{1-\sin\alpha\cos\sigma}{1+\sin\alpha\cos\sigma}\right)^{1/2\sin\alpha}.\,(1-\sin^2\alpha\cos^2\sigma);$$

σ being positive and not greater than $\pi/2$.

Take, *e.g.*,

$$\alpha = 30°.$$

The approximate ellipse is

$$a = 2{\cdot}281, \quad b = 1{\cdot}623,$$

giving

$$\mathrm{R}/\mathrm{R}_0 = 0{\cdot}47, \quad \tfrac{1}{2}\pi - \theta_{\mathrm{A}} = 60°;$$

and the maximum difference in the radii of curvature is at about $\sigma = \pi/4$ and less than 6 per cent. Now write for a second approximation

$$\zeta = \left(\frac{1+\tau}{1-\tau}\right)\left(\frac{1+a\tau}{1-a\tau}\right)^{-1/\sin\alpha} . \, e^{\frac{1}{3}a_3\tau^3+\frac{1}{5}a_3\tau^5},$$

so that

$$\theta = \tfrac{1}{2}\pi - \frac{1}{\sin\alpha}\tan^{-1}(\tan\alpha\sin\sigma) + \tfrac{1}{3}a_3\sin 3\sigma + \tfrac{1}{5}a_3\sin 5\sigma \,;$$

$$\frac{ds}{d\theta} = -\frac{2(1+\cos\sigma)}{\cos\alpha}\left(\frac{1-\sin\alpha\cos\sigma}{1+\sin\alpha\cos\sigma}\right)^{1/2\sin\alpha} .$$

$$(1-\sin^2\alpha\cos^2\sigma)\,\frac{e^{\frac{1}{3}a_3\cos 3\sigma+\frac{1}{5}a_3\cos 5\sigma}}{1-2a_3\cos 4\sigma/\cos\alpha \,.\, (1-\sin^2\alpha\cos^2\sigma)}\,;$$

for σ between 0 and $\pi/2$. Assuming a_3 small, we get the correction factors

$$[1+(\tfrac{8}{15}+2\cos\alpha)a_3] \text{ at } \sigma = 0\,; \quad [1-(4\sqrt{2}/15+\cos\alpha+\sec\alpha)a_3] \text{ at } \sigma = \tfrac{1}{4}\pi\,;$$
$$(1+2a_3\sec\alpha) \text{ at } \sigma = \tfrac{1}{2}\pi\,;$$

in the general case. Divide by the geometric mean of the first and third, and choose a_3, so that the resulting middle correction factor compensates for the difference between the radii of curvature of the barrier and ellipse at $\sigma = \pi/4$. When $\alpha = \pi/6$, the correction factors are

$$(1+2{\cdot}265\,a_3) \text{ at } \sigma = 0\,; \quad (1-2{\cdot}398\,a_3) \text{ at } \sigma = \tfrac{1}{4}\pi\,; \quad (1+2{\cdot}309\,a_3) \text{ at } \sigma = \tfrac{1}{2}\pi\,;$$

so that if we divide by the mean of the first and third, we get

$$(1-0{\cdot}022\,a_3) \text{ at } \sigma = 0\,; \quad (1-4{\cdot}685\,a_3) \text{ at } \sigma = \tfrac{1}{4}\pi\,; \quad (1+0{\cdot}022\,a_3) \text{ at } \sigma = \tfrac{1}{2}\pi.$$

Putting

$$1-4{\cdot}685\,a_3 = 1{\cdot}0577,$$

we deduce

$$a_3 = -0{\cdot}0123.$$

In this case, if we multiply the dimensions of the ellipse already found by 0·972, we have a very good fit, the error being only a fraction of 1 per cent. The ellipse thus found is given by

$$a = 2{\cdot}218, \quad b = 1{\cdot}578,$$

and

$$\mathrm{R}/\mathrm{R}_0 = 0{\cdot}487, \quad \tfrac{1}{2}\pi - \theta_{\mathrm{A}} = 59^\circ\, 51'.$$

The approximation will not be so rapid if α becomes large, but for all values of α the form (9) does give a reasonable first approximation. The case $\alpha = 0$ is that discussed above, § 6. The case $\alpha = \pi/2$ is trivial; it is when $b/a = 0$, so that the ellipse flattens down into a straight line, and there is really no barrier at all. If we take a value of α close to $\pi/2$, say $\sin^{-1} 0{\cdot}999$, we prove that as $b/a \to 0$, $\mathrm{R}/\mathrm{R}_0 \to 0$.

9. In fig. 2 we have given in graphical form the results of calculations

carried out on the lines of the method here sketched. The different curves give the following information :—

I. The value of R/R_0 for b/a lying between 0 and 1·8. For larger values of b/a the labour involved is rather considerable, probably more than is commensurate with the practical interest of the result.

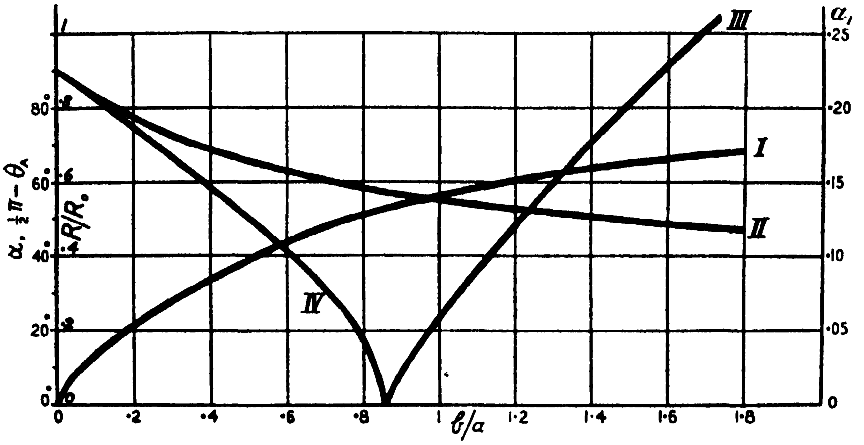


FIG. 2.

II. The value of $\pi/2-\theta_A$ for the same range of b/a.

III. The commencing value of a_1 for b/a greater than 0·858.

IV. The commencing value of α for b/a less than 0·858.

It is obvious that when $b/a = \infty$, $R/R_0 = 1$, $\pi/2-\theta_A = 0$.

It may be of interest to note that the results of our calculations can be summed up in the empirical formula

$$R/R_0 = (b+\tfrac{2}{15}\sqrt{ab})/(a+b) \quad \text{or} \quad R/R_0 = (1+\tfrac{2}{15}\sqrt{x})/(1+x),$$

where x is a/b, the "fineness ratio." This is very nearly correct if $x > 1$, the really useful range for practical purposes. The empirical formula

$$\frac{R}{R_0} = \frac{1}{1+x} + \frac{0{\cdot}185}{x} e^{-1/x}$$

seems to apply to a wider range.

It should now be sufficiently clear that Levi-Civita's transformation enables us to solve problems with curved barriers in a practical manner. If the configurations are unsymmetrical similar methods can be employed—but certain points need special investigation, as will be shown in a further paper.

The Propagation of Earthquake Waves.

By J. H. JEANS, Sec. R.S.

(Received November 8, 1922.)

1. A complete understanding of the propagation of earthquake disturbances demands a solution of the problem of the vibrations of an elastic sphere in which the mutual gravitational attraction of the parts is taken into account, as well as the variation of the physical constants of the earth's material with distance from the centre. The classical solution of Lamb* applied only to a sphere in which gravitation is supposed negligible, and in which the physical constants have uniform values throughout. In a subsequent discussion of certain aspects of the general problem,† I tried to take gravitation into account, but was only able to make progress by introducing the quite artificial assumption of a permanent field of force exactly annulling the gravitational field of the sphere in its undisturbed state. In a valuable critical note on this paper, the late Lord Rayleigh‡ indicated a method which opened up possibilities of progress while avoiding this very artificial restriction. In the present note, I have availed myself of Lord Rayleigh's suggestion, and have treated the problem in certain of its aspects which seem to have a direct seismological importance. A summary and discussion of results will be found at the end (§ 18).

2. We assume the earth to be spherical in shape and to be arranged in spherical layers, so that all the physical constants are functions of a single variable, r, the distance from the centre. In the absence of earthquake disturbances, the stress system at any point is assumed to reduce to a simple hydrostatic pressure, p. This will be connected with the density ρ and the gravitational potential V by the usual equations

$$\frac{\partial p}{\partial x} = \rho \frac{\partial V}{\partial x}, \text{ etc.} \tag{1}$$

in which p, ρ and V are functions of r only.

3. When earthquake disturbances are present, we suppose that the element initially at x, y, z is displaced to $x+\xi, y+\eta, z+\zeta$. If the centre is taken as origin, the radial component of displacement w will be

$$w = \frac{1}{r}(x\xi + y\eta + z\zeta),$$

---

* 'Proc. Lond. Math. Soc.,' vol. 13, p. 189.

† 'Phil. Trans.,' A, vol. 201, p. 157.

‡ 'Roy. Soc. Proc.,' A, vol. 77, p. 486.

and the dilatation Δ is given by

$$\Delta = \frac{\partial \xi}{\partial x}+\frac{\partial \eta}{\partial y}+\frac{\partial \zeta}{\partial z}.$$

The density at x, y, z will be

$$\rho-\frac{\partial}{\partial x}(\rho\xi)-\frac{\partial}{\partial y}(\rho\eta)-\frac{\partial}{\partial z}(\rho\zeta),$$

which, since ρ is assumed to be a function of r only, may be expressed as

$$\rho-\rho\Delta-w\frac{\partial \rho}{\partial r}.$$

Thus the gravitational potential at x, y, z will be V—E, where E is the potential of—

(i) A volume distribution of matter of density $\rho\Delta+w\,(\partial\rho/\partial r)$;

(ii) A surface distribution of matter of density $\rho_0 w$, where ρ_0 is the density at the surface.

It follows that the gravitational potential at $x+\xi$, $y+\eta$, $z+\zeta$ will be

$$V+\xi\frac{\partial V}{\partial x}+\eta\frac{\partial V}{\partial y}+\zeta\frac{\partial V}{\partial z}+\tfrac{1}{2}\xi^2\frac{\partial^2 V}{\partial x^2}+\ldots-E-\xi\frac{\partial E}{\partial x}-\ldots,$$

so that the gravitational force, in the direction of x increasing, at the same point will be

$$\frac{\partial V}{\partial x}+\xi\frac{\partial^2 V}{\partial x^2}+\eta\frac{\partial^2 V}{\partial x\partial y}+\zeta\frac{\partial^2 V}{\partial x\partial z}-\frac{\partial E}{\partial x}.$$

Suppose that, as a result of the displacement, the stress-system at the point initially at x, y, z, but now at $x+\xi$, $y+\eta$, $z+\zeta$, is changed from the assumed hydrostatic pressure,

$$-p,\ -p,\ -p,\ 0,\ 0,\ 0$$

to the more general system

$$-p+T_{xx},\ -p+T_{yy},\ -p+T_{zz},\ T_{yz},\ T_{zx},\ T_{xy}.$$

Then the equations of the element initially at x, y, z, will be

$$\rho\frac{d^2\xi}{dt^2}=-\frac{\partial p}{\partial x}+\frac{\partial T_{xx}}{\partial x}+\frac{\partial T_{xy}}{\partial y}+\frac{\partial T_{xz}}{\partial z}+\rho\left(\frac{\partial V}{\partial x}+\xi\frac{\partial^2 V}{\partial x^2}+\eta\frac{\partial^2 V}{\partial x\partial y}+\zeta\frac{\partial^2 V}{\partial x\partial z}-\frac{\partial E}{\partial x}\right),$$

or, in virtue of equation (1),

$$\rho\frac{d^2\xi}{dt^2}=\frac{\partial T_{xx}}{\partial x}+\frac{\partial T_{xy}}{\partial y}+\frac{\partial T_{xz}}{\partial z}+\rho\left(\xi\frac{\partial^2 V}{\partial x^2}+\eta\frac{\partial^2 V}{\partial x\partial y}+\zeta\frac{\partial^2 V}{\partial x\partial z}-\frac{\partial E}{\partial x}\right). \qquad (2)$$

Following Lord Rayleigh's suggestion, we now assume that the additional stresses T_{xx}, T_{xy}, ... are given by stress-strain relations of the usual type,

$$\left.\begin{aligned}T_{xx}&=\lambda\Delta+2\mu\frac{\partial \xi}{\partial x},\\ T_{yz}&=\mu\left(\frac{\partial \zeta}{\partial y}+\frac{\partial \eta}{\partial z}\right),\ \text{etc.}\end{aligned}\right\} \qquad (3)$$

where λ, μ may be regarded as elastic constants measured with reference to the material as stressed under its own gravitational field in the undisturbed state. They of course reduce to the ordinary elastic constants λ, μ at the earth's surface.

4. From now onwards we shall use spherical polar co-ordinates θ, ϕ, r. We assume the displacement at any point (θ, ϕ, r) to have components u, v, w in the directions of θ, ϕ, r increasing. If G_θ, G_ϕ, G_r are the gravitational components of force in these directions, the equations of motion (2) are found to assume the forms*

$$\rho\left(\frac{\partial^2 u}{\partial t^2}-G_\theta\right)=\frac{1}{r}\frac{\partial T_{\theta\theta}}{\partial\theta}+\frac{1}{r\sin\theta}\frac{\partial T_{\theta\phi}}{\partial\phi}+\frac{3T_{r\theta}+(T_{\theta\theta}-T_{\phi\phi})\cot\theta}{r}+\frac{\partial T_{r\theta}}{\partial r}, \tag{4}$$

$$\rho\left(\frac{\partial^2 v}{\partial t^2}-G_\phi\right)=\frac{1}{r}\frac{\partial T_{\theta\phi}}{\partial\theta}+\frac{1}{r\sin\theta}\frac{\partial T_{\phi\phi}}{\partial\phi}+\frac{3T_{\phi r}+2T_{\theta\phi}\cot\theta}{r}+\frac{\partial T_{\phi r}}{\partial r}, \tag{5}$$

$$\rho\left(\frac{\partial^2 w}{\partial t^2}-G_r\right)=\frac{1}{r}\frac{\partial T_{r\theta}}{\partial\theta}+\frac{1}{r\sin\theta}\frac{\partial T_{\phi r}}{\partial\phi}+\frac{T_{r\theta}\cot\theta+2T_{rr}-T_{\theta\theta}-\partial T_{\phi\phi}}{r}+\frac{\partial T_{rr}}{\partial r}, \tag{6}$$

in which the stress-components are given by

$$T_{\theta\theta}=\lambda\Delta+2\mu\left(\frac{1}{r}\frac{\partial u}{\partial\theta}+\frac{w}{r}\right),$$

$$T_{\phi\phi}=\lambda\Delta+2\mu\left(\frac{1}{r\sin\theta}\frac{\partial v}{\partial\phi}+\frac{u}{r}\cot\theta+\frac{w}{r}\right),$$

$$T_{rr}=\lambda\Delta+2\mu\frac{\partial w}{\partial r},$$

$$T_{\phi r}=\mu\left(\frac{1}{r\sin\theta}\frac{\partial w}{\partial\phi}+\frac{\partial v}{\partial r}-\frac{v}{r}\right),$$

$$T_{r\theta}=\mu\left(\frac{1}{r}\frac{\partial w}{\partial\theta}+\frac{\partial u}{\partial r}-\frac{u}{r}\right),$$

$$T_{\theta\phi}=\mu\left(\frac{1}{r}\frac{\partial v}{\partial\theta}+\frac{1}{r\sin\theta}\frac{\partial u}{\partial\phi}-\frac{v}{r}\cot\theta\right).$$

The boundary conditions require that the normal tension and the two tangential tractions shall vanish at all points on the earth's surface. Analytically, the conditions are that T_{rr}, $T_{\phi r}$ and $T_{r\theta}$ shall all vanish when $r=a$.

From the assumed spherical arrangement of the earth's interior, it follows that V, the gravitational potential in the undisturbed state, must be a function of r only. Hence we readily find, on comparison with equation (2),

---

* Love, 'Theory of Elasticity,' vol. 1, p. 215.

$$G_\theta = -\frac{1}{r}\frac{\partial E}{\partial \theta},$$

$$G_\phi = -\frac{1}{r\sin\theta}\frac{\partial E}{\partial \phi},$$

$$G_r = -\frac{\partial E}{\partial r} + w\frac{\partial^2 V}{\partial r^2}.$$

5. We proceed to substitute these values for the stress-components and the forces in equations (4), (5) and (6). The resulting equations will differ in form from those for a homogeneous sphere, since the elastic-constants λ and μ are now subject to differentiation. But, on account of the assumed spherical arrangement of the earth, λ and μ will be functions of r only, so that the only terms containing differentials of λ and μ will arise from the three last terms, namely, $\partial T_{r\theta}/\partial r$, $\partial T_{\phi r}/\partial r$ and $\partial T_{rr}/\partial r$ in equations (4), (5) and (6).

On dividing throughout by ρ, the equations of motion (4), (5) and (6) assume the forms

$$\frac{\partial^2 u}{\partial t^2} = \frac{\lambda+2\mu}{\rho}\frac{1}{r}\frac{\partial \Delta}{\partial \theta} - \frac{2\mu}{\rho}\frac{1}{r\sin\theta}\frac{\partial \varpi_3}{\partial \phi} + \frac{2\mu}{\rho}\frac{1}{r}\frac{\partial}{\partial r}(r\varpi_2)$$
$$+\frac{1}{\rho}\frac{\partial \mu}{\partial r}\left(\frac{\partial u}{\partial r} + \frac{1}{r}\frac{\partial w}{\partial \theta} - \frac{u}{r}\right) - \frac{1}{r}\frac{\partial E}{\partial \theta}. \qquad (7)$$

$$\frac{\partial^2 v}{\partial t^2} = \frac{\lambda+2\mu}{\rho}\frac{1}{r\sin\theta}\frac{\partial \Delta}{\partial \phi} - \frac{2\mu}{\rho}\frac{1}{r}\frac{\partial}{\partial r}(r\varpi_1) + \frac{2\mu}{\rho}\frac{1}{r}\frac{\partial \varpi_3}{\partial \theta}$$
$$+\frac{1}{\rho}\frac{\partial \mu}{\partial r}\left(\frac{1}{r\sin\theta}\frac{\partial w}{\partial \phi} + \frac{\partial v}{\partial r} - \frac{v}{r}\right) - \frac{1}{r\sin\theta}\frac{\partial E}{\partial \phi}. \qquad (8)$$

$$\frac{\partial^2 w}{\partial t^2} = \frac{\lambda+2\mu}{\rho}\frac{\partial \Delta}{\partial r} + \frac{2\mu}{\rho}\frac{1}{r\sin\theta}\left[\frac{\partial \varpi_1}{\partial \phi} - \frac{\partial}{\partial \theta}(\varpi_2\sin\theta)\right]$$
$$+\frac{1}{\rho}\frac{\partial \lambda}{\partial r}\Delta + \frac{2}{\rho}\frac{\partial \mu}{\partial r}\frac{\partial w}{\partial r} - \frac{\partial E}{\partial r} + w\frac{\partial^2 V}{\partial r^2}. \qquad (9)$$

In these equations ϖ_1, ϖ_2, ϖ_3 are the usual rotations now specified by

$$2\varpi_1 = \frac{1}{r\sin\theta}\frac{\partial w}{\partial \phi} - \frac{1}{r}\frac{\partial}{\partial r}(vr), \qquad (10)$$

$$2\varpi_2 = \frac{1}{r}\frac{\partial}{\partial r}(ur) - \frac{1}{r}\frac{\partial w}{\partial \theta}, \qquad (11)$$

$$2\varpi_3 = \frac{1}{r\sin\theta}\left[\frac{\partial}{\partial \theta}(v\sin\theta) - \frac{\partial u}{\partial \phi}\right], \qquad (12)$$

while the cubical dilatation Δ is given by

$$\Delta = \frac{1}{r^2\sin\theta}\left[\frac{\partial}{\partial \theta}(ur\sin\theta) + \frac{\partial}{\partial \phi}(vr) + \frac{\partial}{\partial r}(wr^2\sin\theta)\right]. \qquad (13)$$

6. Let us introduce D (an operator) and F defined by

$$\frac{\partial^2 \chi}{\partial t^2}-\frac{\mu}{\rho}\frac{1}{r}\frac{\partial^2}{\partial r^2}(\chi r)-\frac{1}{\rho}\frac{\partial \mu}{\partial r}\left(\frac{\partial \chi}{\partial r}-\frac{\chi}{r}\right) \equiv \mathrm{D}\chi, \tag{14}$$

$$\frac{\lambda+2\mu}{\rho}\Delta-\frac{\mu}{\rho}\frac{\partial w}{\partial r}+\frac{1}{\rho}\frac{\partial \mu}{\partial r}w-\mathrm{E} \equiv \mathrm{F}, \tag{15}$$

then the equations (7) and (8) become

$$\mathrm{D}u = -\frac{2\mu}{\rho}\frac{1}{r\sin\theta}\frac{\partial \varpi_3}{\partial \phi}+\frac{1}{r}\frac{\partial \mathrm{F}}{\partial \theta}, \tag{16}$$

$$\mathrm{D}v = \frac{2\mu}{\rho}\frac{1}{r}\frac{\partial \varpi_3}{\partial \theta}+\frac{1}{r\sin\theta}\frac{\partial \mathrm{F}}{\partial \phi}. \tag{17}$$

The elimination of F gives

$$\mathrm{D}\left[\frac{\partial}{\partial \theta}(v\sin\theta)-\frac{\partial u}{\partial \phi}\right] = \frac{2\mu}{\rho}\left[\frac{1}{r}\frac{\partial}{\partial \theta}\left(\sin\theta\,\frac{\partial \varpi_3}{\partial \theta}\right)+\frac{1}{r\sin\theta}\frac{\partial^2 \varpi_3}{\partial \phi^2}\right],$$

or, using relations (12) and (14),

$$\frac{\partial^2 \varpi_3}{\partial t^2} = \frac{\mu}{\rho}\left[\frac{1}{r}\frac{\partial^2}{\partial r^2}(r\varpi_3)+\frac{1}{r^2\sin\theta}\frac{\partial}{\partial \theta}\left(\sin\theta\,\frac{\partial \varpi_3}{\partial \theta}\right)+\frac{1}{r^2\sin^2\theta}\frac{\partial^2 \varpi_3}{\partial \phi^2}\right], \\ +\frac{1}{\rho}\frac{\partial \mu}{\partial r}\left[\frac{\partial \varpi_3}{\partial r}-\frac{\varpi_3}{r}\right],$$

which again can be put in the form

$$\frac{\partial^2 \varpi^3}{\partial t^2} = \frac{\mu}{\rho}\nabla^2\varpi_3+\frac{1}{\rho}\frac{\partial \mu}{\partial r}\left[\frac{\partial \varpi_3}{\partial r}\quad\frac{\varpi_3}{r}\right], \tag{18}$$

an equation involving ϖ_3 only.

7. From the values of ϖ_1, ϖ_2 and Δ given in equations (10), (11) and (13), we obtain

$$2\left[\frac{\partial \varpi_1}{\partial \phi}-\frac{\partial}{\partial \theta}(\varpi_2\sin\theta)\right] = \frac{\sin\theta}{r}\left[r^2\nabla^2 w-\frac{\partial}{\partial r}(r^2\nabla)\right], \tag{19}$$

so that equation (9) can be written in the alternative form

$$\frac{\partial^2 w}{\partial t^2} = \frac{\lambda+2\mu}{\rho}\frac{\partial \Delta}{\partial r}+\frac{\mu}{\rho}\left[\nabla^2 w-\frac{1}{r^2}\frac{\partial}{\partial r}(r^2\Delta)\right]+\frac{1}{\rho}\frac{\partial \lambda}{\partial r}\Delta+\frac{2}{\rho}\frac{\partial \mu}{\partial r}\frac{\partial w}{\partial r}-\frac{\partial \mathrm{E}}{\partial r}+w\frac{\partial^2 \mathrm{V}}{\partial r^2}. \tag{20}$$

We next multiply equations (7), (8) and (9) throughout by $r\sin\theta$, r and $r^2\sin\theta$, operate on the resulting equations with the differential operators

$$\frac{1}{r^2\sin\theta}\frac{\partial}{\partial \theta},\quad \frac{1}{r^2\sin\theta}\frac{\partial}{\partial \phi},\ \text{and}\ \frac{1}{r^2\sin\theta}\frac{\partial}{\partial r},$$

and add corresponding sides. We obtain

$$\begin{aligned}\frac{\partial^2\Delta}{\partial t^2} = \frac{\lambda+2\mu}{\rho}\nabla^2\Delta + \frac{\partial\Delta}{\partial r}\frac{\partial}{\partial r}\left(\frac{\lambda+2\mu}{\rho}\right) + 2\left[\frac{\partial\varpi_1}{\partial\phi} - \frac{\partial}{\partial\theta}(\varpi_2\sin\theta)\right]\frac{1}{r\sin\theta}\frac{\partial}{\partial r}\left(\frac{\mu}{\rho}\right) \\ +\frac{1}{\rho}\frac{\partial\mu}{\partial r}\left[\frac{1}{r\sin\theta}\frac{\partial}{\partial\theta}\left\{\sin\theta\left(\frac{\partial u}{\partial r}+\frac{1}{r}\frac{\partial w}{\partial\theta}-\frac{u}{r}\right)\right\}\right. \\ \left.+\frac{1}{r\sin\theta}\frac{\partial}{\partial\phi}\left\{\frac{1}{r\sin\theta}\frac{\partial w}{\partial\phi}+\frac{\partial v}{\partial r}-\frac{v}{r}\right\}+\frac{2}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial w}{\partial r}\right)\right] \\ +\frac{2}{r^2}\frac{\partial}{\partial r}\left(\frac{r^2}{\rho}\frac{\partial\mu}{\partial r}\right)\frac{\partial w}{\partial r}+\frac{1}{r^2}\frac{\partial}{\partial r}\left(\Delta\frac{r^2}{\rho}\frac{\partial\lambda}{\partial r}\right)-\nabla^2\mathrm{E}+\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2 w\frac{\partial^2\mathrm{V}}{\partial r^2}\right). \quad (21)\end{aligned}$$

The coefficient of $\partial/\partial r\,(\mu/\rho)$ in the first line may be replaced by its value as given by equation (19). The coefficient of $1/\rho\,(\partial\mu/\partial r)$ in the second and third lines may be evaluated by similar methods, and is found to be

$$\frac{\partial\Delta}{\partial r}+\frac{2}{r^2}w+\nabla^2 w.$$

Finally, the value of the term $\nabla^2\mathrm{E}$ is given by Poisson's relation,

$$\nabla^2\mathrm{E} = -4\pi\left(\rho\Delta + w\frac{\partial\rho}{\partial r}\right).$$

Making these substitutions equation (21) assumes the form

$$\begin{aligned}\frac{\partial^2\Delta}{\partial t^2} = \frac{\lambda+2\mu}{\rho}\nabla^2\Delta + \frac{\partial\Delta}{\partial r}\frac{\partial}{\partial r}\left(\frac{\lambda+2\mu}{\rho}\right)+\frac{\partial}{\partial r}\left(\frac{\mu}{\rho}\right)\left[\nabla^2 w - \frac{1}{r^2}\frac{\partial}{\partial r}(r^2\Delta)\right] \\ +\frac{1}{\rho}\frac{\partial\mu}{\partial r}\left(\frac{\partial\Delta}{\partial r}+\frac{2}{r^2}w+\nabla^2 w\right)+2\frac{\partial w}{\partial r}\frac{\partial}{\partial r}\left(\frac{1}{\rho}\frac{\partial\mu}{\partial r}\right)+\frac{\partial\Delta}{\partial r}\frac{1}{\rho}\frac{\partial\lambda}{\partial r} \\ +\frac{2\Delta}{r}\frac{1}{\rho}\frac{\partial\lambda}{\partial r}+\Delta\frac{\partial}{\partial r}\left(\frac{1}{\rho}\frac{\partial\lambda}{\partial r}\right)+\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2 w\frac{\partial^2\mathrm{V}}{\partial r^2}\right)+4\pi\left(\rho\Delta+w\frac{\partial\rho}{\partial r}\right). \quad (22)\end{aligned}$$

8. It has now been seen that the three equations of motion (7), (8) and (9) may be replaced by the three equations (18), (20) and (22). Of these three equations, one involves only the single variable ϖ_3, while the two others involve only the two variables w and Δ. It is natural to regard the general solution of our equations as made up of two parts, one in which w and Δ both vanish, while ϖ_3 satisfies equation (18), and one in which ϖ_3 vanishes, while w and Δ satisfy equations (20) and (22). We shall, of course, not know whether these correspond to physically distinct motions until we have examined the boundary conditions.

9. We first examine the solution in which w and Δ both vanish. Equation (13) gives

$$\frac{\partial}{\partial\theta}(u\sin\theta)+\frac{\partial v}{\partial\phi} = 0,$$

so that we may write

$$u \sin\theta = \frac{\partial S}{\partial\phi}, \quad v = -\frac{\partial S}{\partial\theta}, \tag{23}$$

where S is a function which remains to be determined.

When Δ and w both vanish, $E = 0$, so that $F = 0$ from equation (15). The third equation of motion (equation (20)) is automatically satisfied, while the first two equations of motion (equations (16) and (17)) assume the forms

$$D(u \sin\theta) = -\frac{2\mu}{\rho r}\frac{\partial\varpi_3}{\partial\phi}, \tag{24}$$

$$Dv = \frac{2\mu}{\rho r}\frac{\partial\varpi_3}{\partial\theta}. \tag{25}$$

On inserting the values for $u \sin\theta$ and v given by equations (23), these equations are seen to have the common integral

$$DS = -\frac{2\mu}{\rho r}\varpi_3. \tag{26}$$

It is not necessary to add a function of r, since the addition of such a function would not affect the values of the displacements u, v, w, as determined by equations (23). On inserting the value of ϖ_3 from equation (12), this becomes

$$DS = \frac{\mu}{\rho}\left[\frac{1}{r^2 \sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial S}{\partial\theta}\right) + \frac{1}{r^2 \sin^2\theta}\frac{\partial^2 S}{\partial\phi^2}\right],$$

or, restoring the value of D,

$$\frac{\partial^2 S}{\partial t^2} = \frac{\mu}{\rho}\nabla^2 S + \frac{1}{\rho}\frac{\partial\mu}{\partial r}\left[\frac{\partial S}{\partial r} - \frac{S}{r}\right]. \tag{27}$$

If S is any solution of this equation, a displacement specified by equations (23) will satisfy all the equations of motion.

We turn to a consideration of the boundary conditions. Since our solution makes $w = 0$ and $\Delta = 0$, the first boundary condition ($T_{rr} = 0$ on p. 556) is already satisfied, while the two remaining conditions ($T_{\phi r} = 0$ and $T_{r\theta} = 0$) assume the forms

$$\frac{\partial}{\partial\theta}\left(\frac{\partial S}{\partial r} - \frac{S}{r}\right) = 0,$$

$$\frac{1}{\sin\theta}\frac{\partial}{\partial\phi}\left(\frac{\partial S}{\partial r} - \frac{S}{r}\right) = 0,$$

and these are simultaneously satisfied if

$$\frac{\partial S}{\partial r} - \frac{S}{r} = 0 \text{ at } r = a. \tag{28}$$

Thus a solution of the whole problem is given by equations (23) if S is such as to satisfy equation (27) throughout the earth's interior, and equation (28) at the earth's surface.

10. We consider next solutions in which ϖ_3 vanishes, but both w and Δ exist. When $\varpi_3 = 0$, equation (12) gives

$$\frac{\partial}{\partial\theta}(v\sin\theta) = \frac{\partial u}{\partial\phi}, \tag{29}$$

and it is natural to assume a solution in the form

$$\left.\begin{aligned} u &= \frac{1}{r}\frac{\partial \mathrm{P}}{\partial\theta} \\ v &= \frac{1}{r\sin\theta}\frac{\partial \mathrm{P}}{\partial\phi} \\ w &= \frac{\partial \mathrm{P}}{\partial r} + w' \end{aligned}\right\}, \tag{30}$$

in which P is a function which remains to be determined. This solution is the most general solution which is consistent with relation (29), provided w' is regarded as quite general.

With these values for the displacements, the motion is obviously "irrotational" except for the terms in w', and on calculating the actual rotations, we find

$$\left.\begin{aligned} 2\varpi_1 &= \frac{1}{r\sin\theta}\frac{\partial w'}{\partial\phi} \\ 2\varpi_2 &= -\frac{1}{r}\frac{\partial w'}{\partial\theta} \\ 2\varpi_3 &= 0 \end{aligned}\right\}. \tag{31}$$

The cubical dilatation Δ is found to be given by

$$\Delta = \nabla^2\mathrm{P} + \frac{1}{r^2}\frac{\partial}{\partial r}(r^2 w'). \tag{32}$$

The first two equations of motion (16) and (17) now assume the forms—

$$\mathrm{D}\left(\frac{1}{r}\frac{\partial \mathrm{P}}{\partial\theta}\right) = \frac{1}{r}\frac{\partial \mathrm{F}}{\partial\theta},$$

$$\mathrm{D}\left(\frac{1}{r\sin\theta}\frac{\partial \mathrm{P}}{\partial\phi}\right) = \frac{1}{r\sin\theta}\frac{\partial \mathrm{F}}{\partial\phi},$$

and these have the common integral

$$\mathrm{D}\left(\frac{\mathrm{P}}{r}\right) = \frac{\mathrm{F}}{r}. \tag{33}$$

On substituting for D (equation (14)), F (equation (15)), Δ and w, and simplifying, this becomes

$$\frac{d^2\mathrm{P}}{dt^2} = \frac{\lambda+2\mu}{\rho}\nabla^2\mathrm{P} - \frac{2}{\rho}\frac{\partial\mu}{\partial r}\frac{\mathrm{P}}{r} + \frac{\lambda+2\mu}{\rho}\frac{1}{r^2}\frac{\partial}{\partial r}(r^2 w') - \frac{\mu}{\rho}\frac{\partial w'}{\partial r} + \frac{1}{\rho}\frac{\partial\mu}{\partial r}w' - \mathrm{E}. \tag{34}$$

The third equation (9) assumes the form

$$\frac{d^2}{dt^2}\left(\frac{\partial P}{\partial r}\right)+\frac{d^2w'}{dt^2} = \frac{\lambda+2\mu}{\rho}\frac{\partial}{\partial r}(\nabla^2 P)+\frac{\mu}{\rho}\nabla^2 w'+\frac{\lambda+2\mu}{\rho}\frac{\partial}{\partial r}\left[\frac{1}{r^2}\frac{\partial}{\partial r}(r^2 w')\right]$$

$$-\frac{\mu}{\rho}\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial w'}{\partial r}\right)+\frac{1}{\rho}\frac{\partial\lambda}{\partial r}\left[\nabla^2 P+\frac{1}{r^2}\frac{\partial}{\partial r}(r^2 w')\right]$$

$$+\frac{2}{\rho}\frac{\partial\mu}{\partial r}\left(\frac{\partial^2 P}{\partial r^2}+\frac{\partial w'}{\partial r}\right)-\frac{\partial E}{\partial r}+\left(\frac{\partial P}{\partial r}+w'\right)\frac{\partial^2 V}{\partial r^2}. \qquad (35)$$

Thus if P and w' are chosen so as to satisfy these equations, solution (30) will satisfy all the equations of motion.

We consider the boundary conditions. The two equations expressing that the tangential tractions shall vanish ($T_{\phi r} = 0$, $T_{r\theta} = 0$) assume the forms—

$$\frac{1}{r\sin\theta}\frac{\partial w}{\partial\phi}+\frac{1}{r\sin\theta}\frac{\partial}{\partial\phi}\left(\frac{\partial P}{\partial r}-\frac{2P}{r}\right) = 0,$$

$$\frac{1}{r}\frac{\partial w}{\partial\theta}+\frac{1}{r}\frac{\partial}{\partial\theta}\left(\frac{\partial P}{\partial r}-\frac{2P}{r}\right) = 0,$$

and these are both satisfied if

$$w+\frac{\partial P}{\partial r}-\frac{2P}{r} = 0 \text{ at } r = a,$$

or, what is the same thing, if

$$w'+2\left(\frac{\partial P}{\partial r}-\frac{P}{r}\right) = 0 \text{ at } r = a. \qquad (36)$$

The third boundary condition ($T_{rr} = 0$) becomes

$$\lambda\left(\nabla^2 P+\frac{1}{r^2}\frac{\partial}{\partial r}(r^2 w')\right)+2\mu\left(\frac{\partial^2 P}{\partial r^2}+\frac{\partial w'}{\partial r}\right), \qquad (37)$$

and since the boundary conditions involve only P and w', it is clear that by a suitable adjustment of the constants, solutions of the type of (30) can be found to satisfy all the conditions of the problem.

11. When we restrict our problem to the free vibrations of a homogeneous non-gravitating sphere, the two types of solution we have found reduce to the two types of free vibrations which Lamb describes as a vibration of the first class and second class respectively. But when we restrict our problem in a different way, namely, to the propagation of shocks in a natural earth, then the two types of solution correspond respectively to what seismologists describe as secondary and primary waves.

We shall adopt and extend the seismological nomenclature and describe our two types of motion as secondary and primary disturbances respectively. Our equations show that these correspond to separate groups of free vibrations of the earth.

To search for free vibrations, we assume a time factor, $\cos pt$, and replace d^2/dt^2 in our equations by $-p^2$. It then appears that the resulting equations can be solved by assuming P and w' for primary vibrations, or S for secondary vibrations, to be proportional to a surface harmonic, $S_n(\theta, \phi)$. For Δ and w then become proportional to $S_n(\theta, \phi)$, and therefore also E. Thus the free vibrations are deducible from functions of the types

$$\left.\begin{aligned} P &= F_1(r)S_n(\theta, \phi)\cos pt \\ w' &= f_1(r)S_n(\theta, \phi)\cos pt \end{aligned}\right\}\text{(Primary vibrations).}$$

$$S = F_2(r)S_n(\theta, \phi)\cos pt \quad \text{(Secondary vibrations).}$$

12. This is probably as far as we can go with profit in our discussion of the general problem, but farther progress, in one direction at least, is made possible by a simplification which does not seriously mar the seismological value of our result. An earthquake shock consists essentially of the propagation of a disturbance which originates in a small volume of the earth's interior. The volume initially affected is so small that throughout it the range of values in r is small compared with the earth's radius and the range of values in θ, ϕ is small compared with 4π, the solid angle of the earth's surface. It results that when the initial disturbance is analysed into free vibrations a large part of it goes into vibrations in which the displacements change rapidly for small changes, both in r and in θ, ϕ. In other words a natural earthquake shock can be regarded as being made up in great part of free vibrations in which the order, n, of the surface harmonic is large, and in which the functions $F_1(r)$, $f_1(r)$ and $F_2(r)$ vary rapidly with r.

We shall now introduce the simplification of limiting ourselves to vibrations of the type just described. Its effect may be described briefly, although not very accurately, as the neglect of all earthquake waves except those of wave-length short in comparison with the earth's radius.

13. Let us consider the "secondary" disturbance first. This is specified by a single function, S, which satisfies equation (27) throughout the earth's interior and satisfies equation (28) at its surface. On assuming S to be proportional to $S_n(\theta, \phi)\cos pt$, equation (27) reduces to

$$p^2S + \frac{\mu}{\rho}\left[\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial S}{\partial r}\right) - \frac{n(n+1)}{r^2}S\right] + \frac{1}{\rho}\frac{\partial \mu}{\partial r}\left[\frac{\partial S}{\partial r} - \frac{S}{r}\right] = 0. \qquad (38)$$

If we limit ourselves to a small range of values of r, we may, as a consequence of the assumed rapid variation of S with r, neglect any term in S in comparison with a similar term containing a differential of S of higher order with respect to r; we may, in fact, treat $\partial/\partial r$ as a large multiplier. But we must not carry the procedure beyond a small range of values of r since although $\partial P/\partial r$ may be very large in comparison with P/r, it by no means follows that $\int(\partial P/\partial r)$ will be large in comparison with $\int(P/r)\,dr$ when the integration extends over a large range.

Thus over any small range of values of r, equation (38) may be replaced by

$$p^2 S+\frac{\mu}{\rho}\left[\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial S}{\partial r}\right)-\frac{n(n+1)}{r^2}S\right]=0, \tag{39}$$

the last term in this equation being retained because n is supposed to be a large quantity. The solution of this equation is

$$S=r^{-\frac{1}{2}}\{AJ_{n+\frac{1}{2}}[p\sqrt{(\rho/\mu)}\,r]+BJ_{-(n+\frac{1}{2})}[p\sqrt{(\rho/\mu)}\,r]\}\,S_n(\theta,\phi)\cos pt, \tag{40}$$

where J denotes the usual Bessel function. This not only gives the solution of equation (38) through any small range of values of r, but will give the solution throughout the whole range if A and B are regarded as slowly changing qualities. The condition of finiteness at the origin requires that $B=0$ when $r=0$.

Whatever the value of n, it is known that $J_n(x)$ increases steadily with x from $x=0$ to $x=n$. Where n is large it approximates at $x=n$ to the first term of Debye's formula for $J_n(n)$, namely,

$$J_n(n)=\frac{6^{\frac{1}{3}}\Gamma(\frac{1}{3})}{2\pi\,3^{\frac{1}{2}}}\,n^{-\frac{1}{3}}, \tag{41}$$

while when x is substantially greater than n, it is given by the first term of Debye's asymptotic expansion

$$J_n(n\sec\beta)\sim\left(\frac{2}{n\pi\tan\beta}\right)^{\frac{1}{2}}\cos(n\tan\beta-n\beta-\tfrac{1}{4}\pi). \tag{42}$$

Regarding n as a large quantity, we see that $J_n(x)$ is of the order of $n^{-\frac{1}{3}}$ when $x=n$, but rapidly becomes of the order of $n^{-\frac{1}{2}}$ as x exceeds this value. Thus $J_n(x)$ reaches its maximum value in the neighbourhood of $x=n$. When n is large $J_n(x)$ is inappreciable until x approaches the value $x=n$; it here has a positive value, but as soon as x exceeds n the value of $J_n(x)$ oscillates rapidly from positive to negative, ultimately becoming small in comparison with its value at $x=n$.

Within the range of values of r which is close to $r=a$, the value of S may be supposed given by equation (40) in which ρ and μ are assigned the values they have at the surface. Physical solutions will be given by values of S such that (*cf.* equation (28)),

$$\frac{\partial}{\partial r}\left(\frac{S}{r}\right)=0 \text{ at } r=a, \tag{43}$$

while the complete solution for S must make S finite at $r=0$.

Putting $B=0$ in equation (40) we obtain

$$S/r=Ar^{-\frac{3}{2}}J_{n+\frac{1}{2}}[p\sqrt{(\rho/\mu)}\,r], \tag{44}$$

a special solution in which the existence of S is limited to the surface layer near $r = a$, if we have, approximately,

$$p\sqrt{(\rho/\mu)}\, a = n+\tfrac{1}{2}, \tag{45}$$

and so which makes S finite (being zero) at $r = 0$. To satisfy the boundary condition (43) the expression on the right-hand of equation (44) must be a maximum or a minimum at $r = a$. From what has been said, it is clear that the lowest value of p for which this condition is satisfied is very close to the value determined by equation (45), while as p increases above this value other solutions follow in rapid succession.

14. The analysis in terms of Bessel functions is suited for direct comparison with Lamb's discussion of the vibrations of a homogeneous non-gravitating sphere. But a simpler, and at the same time more accurate, method is available for the special problem we have in hand. Again treating $\partial/\partial r$ as a large quantity, it appears that equation (39) may be written in the still simpler form

$$p^2 S + \frac{\mu}{\rho}\left[\frac{\partial^2 S}{\partial r^2} - \frac{n(n+1)}{r^2} S\right] = 0, \tag{46}$$

or

$$\frac{\partial^2 S}{\partial r^2} = n(n+1)\left(\frac{1}{r^2} - \frac{1}{r_0^2}\right) S,$$

where

$$r_0^2 = \frac{n(n+1)}{p^2}\frac{\mu}{\rho}. \tag{47}$$

The solution throughout any small range of values of r in the neighbourhood of $r = b$ is, when $b > r_0$,

$$S = A\cos\left[r\left\{n(n+1)\left(\frac{1}{r_0^2} - \frac{1}{b^2}\right)\right\}^{\frac{1}{2}} - \epsilon\right], \tag{48}$$

while when $b < r_0$, it is

$$S = B\exp\left[-r\left\{n(n+1)\left(\frac{1}{b^2} - \frac{1}{r_0^2}\right)\right\}^{\frac{1}{2}}\right], \tag{49}$$

to which may be added a complementary solution containing the exponential with a positive index, which must ultimately be rejected, so as to make S finite at $r = 0$.

Clearly the first roots of $\frac{\partial}{\partial r}\left(\frac{S}{r}\right) = 0$ will occur close to $r_0 = b$, so that the vibrations of lowest frequency, for which this condition must be satisfied at the earth's surface, are obtained by putting $r_0 = a$, or, what is the same thing,

$$p = \frac{1}{a}\sqrt{\left[n(n+1)\frac{\mu}{\rho}\right]}, \tag{50}$$

a result agreeing with that of equation (45), which was obtained by Bessel function analysis. It is readily found that the succession of roots of p is given by putting $s = 0, 1, 2 \ldots$ in the formula

$$p = \frac{1}{a}\sqrt{\left[\{n(n+1)+(s\pi+\eta)^2\}\frac{\mu}{\rho}\right]}, \tag{51}$$

where η is less than π. This is true so long as s is small in comparison with n, and for such values of s the vibrations are confined to a thin skin on the earth's surface. When s becomes comparable with n, the vibrations extend a perceptible distance into the earth's interior, and formula (51) begins to fail.

15. The high-frequency vibrations we have had under consideration have, so far, been regarded merely as free vibrations each endowed with the energy allotted to it by the original earthquake shock, but we shall now find that they may be interpreted in a more physical manner. Consider, for simplicity, the case in which the initial shock occurred at the point $\theta = 0$ on the earth's surface and spread out symmetrically in all directions. Clearly our surface harmonic $S_n(\theta, \phi)$ may be replaced by the zonal harmonic P_n. This satisfies

$$\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial P_n}{\partial\theta}\right)-n(n+1)P_n = 0,$$

but the procedure already followed enables us to replace this, throughout any small range of values of θ for which $\sin\theta$ is not very small, by

$$\frac{\partial^2 P_n}{\partial\theta^2}-n(n+1)P_n = 0, \tag{52}$$

of which the solutions are

$$\cos[n(n+1)]^{\frac{1}{2}}\theta \text{ and } \sin[n(n+1)]^{\frac{1}{2}}\theta.$$

Combining this with time factors $\cos pt$ and $\sin pt$, it appears that the solutions for S may be supposed proportional to the function of r already discussed and to

$$\cos\{[n(n+1)]^{\frac{1}{2}}\theta \pm pt\}, \tag{53}$$

an expression which represents waves travelling along the meridians $\phi =$ cons., with a uniform surface velocity, v, equal to $ap[n(n+1)]^{-\frac{1}{2}}$. On using the value for p given by formula (51), we have

$$v^2 = \frac{\mu}{\rho}\left[1+\frac{(s\pi+\eta)^2}{n(n+1)}\right]. \tag{54}$$

For pure surface-waves, s has values small compared with n, so that the formula reduces to

$$v^2 = \mu/\rho, \tag{55}$$

giving a value of v which is independent both of s and n.

In the special case in which $\sin\theta$ is small, we may replace it by θ (or $\pi-\theta$), and obtain, in place of equation (52),

$$\frac{\partial^2}{\partial\theta^2}(\mathrm{P}_n\sqrt{\theta})-(n+\tfrac{1}{2})^2\,\mathrm{P}_n\sqrt{\theta}=0,$$

and in place of formula (53),

$$\theta^{-\frac{1}{2}}\cos\{(n+\tfrac{1}{2})\theta\pm pt\} \tag{56}$$

(or a corresponding expression with $\pi-\theta$ replacing θ).

This represents waves travelling with a velocity given by

$$v^2=\frac{\mu}{\rho}\left[1+\frac{(s\pi+\eta)^2-\frac{1}{4}}{(n+\frac{1}{2})^2}\right],$$

a velocity which is indistinguishable from that given by equation (54) down to terms of the order of $1/n^2$.

Thus we see that for all values of n which are large and for all values of s which are small in comparison with n, there are systems of free-vibrations of the earth which may be regarded as surface waves travelling round the earth's surface with the uniform velocity $\sqrt{(\mu/\rho)}$. The point of our somewhat complicated analysis has not, of course, been to show that surface waves of the secondary type travel with a velocity $\sqrt{(\mu/\rho)}$ — this was known already. The essential point is that these surface waves represent a complete system of free vibrations, and so travel round and round the globe without any loss of energy except that due to dissipative forces such as viscosity; there is no tendency for these waves to diffuse into the earth's interior, and, in the neighbourhood of the points $\theta=0$ and $\theta=\pi$ equation (56) shows that the amplitude of these waves increases without limit, being proportional respectively to $\theta^{-\frac{1}{2}}$ and $(\pi-\theta)^{\frac{1}{2}}$, a result which is also obvious from considerations of energy.

16. We may treat the primary disturbance in the same way. Limiting ourselves to waves of short wave-length, we neglect in equations (34) and (35) all terms except those differentiated the maximum number of times with respect to r, and find that these equations assume the forms

$$\frac{d^2\mathrm{P}}{dt^2}=\frac{\lambda+2\mu}{\rho}\nabla^2\mathrm{P}+\frac{\lambda+\mu}{\rho}\frac{\partial w'}{\partial r}, \tag{57}$$

$$\frac{d^2}{dt^2}\left(\frac{\partial\mathrm{P}}{\partial r}\right)+\frac{d^2w'}{dt^2}=\frac{\lambda+2\mu}{\rho}\frac{\partial}{\partial r}(\nabla^2\mathrm{P})+\frac{\mu}{\rho}\nabla^2w'+\frac{\lambda+\mu}{\rho}\frac{\partial^2w'}{\partial r^2}. \tag{58}$$

Our present procedure of dropping all differential coefficients except those of highest order results in $\partial/\partial r$ being commutative with all functions of r. Remembering this, differentiating both sides of equation (57) with respect to r, and subtracting from corresponding sides of equation (58), we obtain

$$\frac{d^2w'}{dt^2}=\frac{\mu}{\rho}\nabla^2w', \tag{59}$$

so that w' satisfies the same equation as S. When w' is proportional to $S_n(\theta, \phi)\cos pt$, this reduces to

$$p^2 w' + \frac{\mu}{\rho}\left[\frac{\partial^2 w'}{\partial r^2} - \frac{n(n+1)}{r^2} w'\right] = 0. \tag{60}$$

The general solution of equation (57), which makes P proportional to $S_n(\theta, \phi)\cos pt$, is readily found to be

$$P = P_0 + \frac{\mu}{\rho p^2}\frac{\partial w'}{\partial r}, \tag{61}$$

where P_0 is the most general solution of

$$p^2 P_0 + \frac{\lambda + 2\mu}{\rho}\left[\frac{\partial^2 P_0}{\partial r^2} - \frac{n(n+1)}{r^2} P_0\right] = 0. \tag{62}$$

The first boundary condition (equation (36)) reduces to

$$w' = -2\frac{\partial P}{\partial r} = -2\frac{\partial P_0}{\partial r} - \frac{2\mu}{\rho p^2}\frac{\partial^2 w'}{\partial r^2}$$

at $r = a$. Using equation (60), this becomes

$$\frac{\partial P_0}{\partial r} = -\left[\frac{\mu}{\rho p^2}\frac{n(n+1)}{a^2} - \tfrac{1}{2}\right] w'. \tag{63}$$

The second boundary condition (equation (37)), becomes

$$(\lambda + 2\mu)\left[\nabla^2 P + \frac{\partial w'}{\partial r}\right] + 2\mu\frac{n(n+1)}{a^2} P = 0 \text{ at } r = a,$$

and this, by use of equation (57), may be expressed as

$$\frac{\partial w'}{\partial r} = P\left(p^2\frac{\rho}{\mu} - \frac{2n(n+1)}{a^2}\right),$$

or again, using equation (61),

$$P_0\left(p^2\frac{\rho}{\mu} - \frac{2n(n+1)}{a^2}\right) = \frac{2n(n+1)}{a^2}\left(\frac{\mu}{\rho p^2}\right)\frac{\partial w'}{\partial r}. \tag{64}$$

From the two boundary-conditions (63) and (64), we obtain

$$\frac{n(n+1)}{a^2}\left(\frac{1}{P_0}\frac{\partial P_0}{\partial r}\right)\left(\frac{1}{w_0'}\frac{\partial w_0'}{\partial r}\right) = \left[\frac{n(n+1)}{a^2} - \tfrac{1}{2}p^2\frac{\rho}{\mu}\right]^2. \tag{65}$$

17. All possibilities are included in the three cases—

(i) $n(n+1) > p^2 a^2 \dfrac{\rho}{\mu} > p^2 a^2 \dfrac{\rho}{\lambda + 2\mu}$,

(ii) $p^2 a^2 \dfrac{\rho}{\mu} > n(n+1) > p^2 a^2 \dfrac{\rho}{\lambda + 2\mu}$,

(iii) $p^2 a^2 \dfrac{\rho}{\mu} > p^2 a^2 \dfrac{\rho}{\lambda + 2\mu} > n(n+1)$.

We shall discuss these three cases in turn.

Case 1.—Following the method of § 15, we may assume that in the neighbourhood of $r = a$ the solutions of equations (60) and (62) are

$$w' = Ae^{-\alpha r}, \quad P_0 = Be^{-\beta r}, \tag{66}$$

where $$\alpha^2 = \frac{n(n+1)}{a^2} - p^2\frac{\rho}{\mu}, \quad \beta^2 = \frac{n(n+1)}{a^2} - p^2\frac{\rho}{\lambda+2\mu}.$$

With this solution we have

$$\frac{1}{w'}\frac{\partial w'}{\partial r} = -\alpha \quad \text{and} \quad \frac{1}{P_0}\frac{\partial P_0}{\partial r} = -\beta,$$

so that equation (65) becomes

$$\alpha\beta\frac{n(n+1)}{a^2} = \left[\frac{n(n+1)}{a^2} - \tfrac{1}{2}p^2\frac{\rho}{\mu}\right]^2, \tag{67}$$

and if this equation is satisfied we can, of course, arrange the ratio A/B so that the two boundary conditions shall be separately satisfied.

Squaring both sides of equation (67), substituting for α^2 and β^2, and simplifying, we obtain

$$(\lambda+2\mu)\theta^3 - 8(\lambda+2\mu)\theta^2 + 8(3\lambda+4\mu)\theta - 16(\lambda+\mu) = 0, \tag{68}$$

where $$p^2 = \frac{n(n+1)}{a^2}\frac{\mu}{\rho}\theta.$$

The surface disturbances we have under consideration are now seen to be identical with the "waves propagated along the plane surface of an elastic solid" discussed by Lord Rayleigh in 1885,* our equation (68) being identical with his equation (24). Our analysis has made it clear that these waves can be propagated around the curved surface of the earth without any escape of energy by diffusion into the interior.

Equation (68) admits of numerical solution when the ratio of λ to μ is known. A velocity of propagation equal to $\sqrt{(\mu/\rho)}$ would require $\theta = 1$, and this is a root of equation (68) only when $\lambda = -2\mu$, a value which is impossible for a natural body. Hence it is readily seen that, whatever the value of λ/μ for the earth, the velocity of propagation is necessarily less than $\sqrt{(\mu/\rho)}$. The following Table gives solutions of equation (68) for various ratios of λ/μ, the values given being those of $\sqrt{\theta}$, the ratio of the velocity of propagation to $\sqrt{(\mu/\rho)}$. The values for $\lambda/\mu = \infty, 1, 0, -\frac{2}{3}$ are due to Lord Rayleigh.

$\lambda/\mu =$	∞	$1\frac{1}{8}$	1	$\frac{3}{4}$	0	$-\frac{2}{3}$	-1
$\sqrt{\theta} =$	0·9554	0·9218	0·9194	0·9135	0·8741	0·6896	0·0

For the earth, λ/μ is probably about $1\frac{1}{8}$, and near this value $\sqrt{\theta}$ is not very sensitive to changes in λ/μ. Thus, without knowing the value of

---

* 'Proc. Lond. Math. Soc.,' vol. 17, p. 4 (1885).

λ/μ with great accuracy, we may assert that the velocity of propagation of Rayleigh waves over the earth's surface is about $0{\cdot}92\sqrt{(\mu/\rho)}$.

Case 2.—When $n(n+1) < p^2a^2\rho/\mu$, the value of w' given by equation (66) must be replaced by

$$w' = \mathrm{A}\cos(\alpha' r - \epsilon) \text{ with } \alpha'^2 = p^2\frac{\rho}{\mu} - \frac{n(n+1)}{a^2}. \tag{69}$$

The value of $\frac{1}{w'}\frac{\partial w'}{\partial r}$ is now $-\alpha'\tan(\alpha' r - \epsilon)$, so that equation (67) must be replaced by

$$\alpha'\beta\tan(\alpha' a - \epsilon) = \left[\frac{n(n+1)}{a^2} - \tfrac{1}{2}p^2\frac{\rho}{\mu}\right]^2. \tag{70}$$

The left-hand member of this equation passes through the whole range of values from $-\infty$, through zero, to $+\infty$ each time that $\alpha' a$ increases by π. Thus it is readily seen that there are a very great number of roots of equation (70) for values of p close up to that given by

$$n(n+1) = p^2a^2\rho/\mu,$$

and these correspond to a system of purely surface waves propagated with a uniform velocity $\sqrt{(\mu/\rho)}$.

Case 3.—Following the same procedure we find that equation (70) must be replaced by

$$\alpha'\beta'\tan(\alpha' a - \epsilon)\tan(\beta' a - \eta) = \left[\frac{n(n+1)}{a^2} - \tfrac{1}{2}p^2\frac{\rho}{\mu}\right]^2, \tag{71}$$

where η is a constant similar to the ϵ introduced in case (2) and

$$\beta'^2 = p^2\frac{\rho}{\lambda + 2\mu} - \frac{n(n+1)}{a^2}.$$

The equation may also be written in the form

$$\alpha'\beta'\tan(\alpha' a - \epsilon) = \left[\frac{n(n+1)}{a^2} - \tfrac{1}{2}p^2\frac{\rho}{\mu}\right]^2\tan\left(\beta' a - \eta + \frac{\pi}{2}\right). \tag{72}$$

As p varies, both sides of this equation repeatedly cover the range of values from $-\infty$ through 0 to $+\infty$, whence it is clear, as previously, that successive roots of the equation are separated only by infinitesimal differences in p. As soon as p exceeds by ever so little the value given by

$$p^2a^2\rho/(\lambda + 2\mu) = n(n+1),$$

there are a great number of roots representing free vibrations of approximately equal periods. These correspond to waves propagated over the earth's surface with velocity $\sqrt{[(\lambda + 2\mu)/\rho]}$.

There is an essential difference between this case and cases 1 and 2 previously considered. In each of the former cases the limiting solution made both P_0 and w' (and hence also Δ and w) vanish throughout the earth, except for a thin surface skin. The present solution makes P_0 so vanish, but w' does not vanish in the earth's interior. On putting $P_0 = 0$ throughout the earth's interior, however, we have

$$P = \frac{\mu}{\rho p^2} \frac{\partial w'}{\partial r},$$

whence we find

$$w = \frac{\mu n(n+1)}{\rho a^2 p^2} w' = \frac{\mu}{\lambda + 2\mu} w',$$

together with $\Delta = 0$. Thus we may regard the motion as consisting of a series of waves of dilatation and compression, only skin-deep, travelling over the earth's surface with the velocity already mentioned, while the earth's interior adjusts itself to the motion by a series of displacements which involve no compression. This concludes our mathematical discussion; we now proceed to summarise and discuss our results.

Discussion and Conclusion.

18. We have regarded the earth as a gravitating compressible sphere, arranged in concentric spherical layers. Assuming that the stress-strain relations in the earth's interior are of the type suggested by the late Lord Rayleigh, we have studied the free vibrations of this sphere, especially in reference to their interpretation as waves spreading out from a centre at or near the earth's surface.

We have found first that all possible free vibrations fall into two distinct classes, corresponding exactly to the primary and secondary waves of seismology.

The secondary class are the simplest, the tangential and radial displacements being of the types

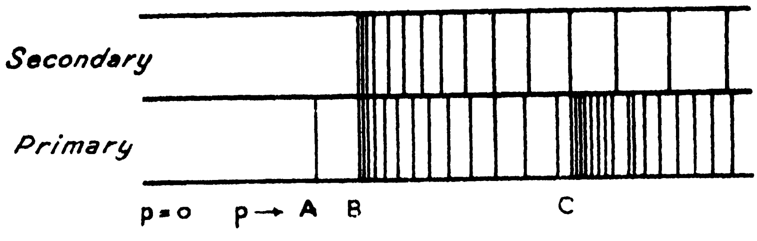
$$\frac{1}{\sin\theta}\frac{\partial S}{\partial \phi}, \quad -\frac{\partial S}{\partial \theta}, \quad 0,$$

where S is a function satisfying a simple differential equation. The primary class show displacements of the type

$$\frac{\partial P}{\partial \theta}, \quad \frac{\partial P}{\sin\theta\, \partial \phi}, \quad \frac{\partial P}{\partial r} + w',$$

where P and w' satisfy rather less simple differential equations. For a single primary free vibration P and w' are proportional to the same surface harmonic $S_n(\theta, \phi)$; for a single secondary free vibration S is also proportional to a surface harmonic.

We have studied the frequencies of these vibrations. Corresponding to a single surface harmonic of order n, the earth's "spectrum" is found to be of the type shown in the accompanying figure.



The point A corresponds to a solitary surface wave of the type discussed by Lord Rayleigh, travelling round and round the earth's surface. Its velocity is about $0{\cdot}92\sqrt{(\mu/\rho)}$, and so is independent of n. The lines at B represent great numbers of surface waves of a different type travelling round and round with uniform velocity $\sqrt{(\mu/\rho)}$. As we pass from B to higher frequencies the waves travel over the surface with higher velocity, but are no longer confined entirely to the surface; there is diffusion into the earth's interior. Again, the lines at C represent great numbers of a still different type, travelling round and round with the uniform velocity $\sqrt{[(\lambda+2\mu)/\rho]}$. In these waves the dilatation is wholly confined to the surface, but is accompanied by a non-dilatational motion of adjustment in the earth's interior. To the right of C we have waves in which both dilatation and rotation are spread through the earth's interior. There is a similar spectrum for each order n of surface harmonic, and the critical velocities represented at A, B, and C, being independent of n, are the same for all.

Our method of regarding these waves as compounded of free vibrations has shown that the waves represented at A, B, and C travel round and round without any escape of energy into the earth's interior. They may, in fact, be regarded as being propagated in two dimensions only. Lord Rayleigh has already remarked on this property as applying to his surface wave, represented at A, and has suggested that in consequence such waves may be of great importance in earthquakes and collisions. But it is now clear that these Rayleigh waves form only an insignificant fraction of the total of waves possessing this property, so that their seismological importance will be infinitesimal in comparison with that of the waves represented at B and C.

Let t_1, t_2 be the times required for the waves represented at C and B respectively to travel completely round the earth. Let an earthquake occur at a point O on the earth's surface at time $t = 0$, and result in the emission from O of surface waves of these two types. After a time $\frac{1}{2}t_1$, the C-waves, travelling along all the different meridians through O, will come to a focus at O′, the antipodes of O, and a seismograph situated close to O′ might record the passage of these waves.

After a time, t_1, the C-waves will again be focussed at O, the site of the original earthquake. It is possible that here a natural seismograph may be provided by the structure of the earth, already shaken and weakened by the first shock, and that the simultaneous arrival of the army of C-waves may be recorded as a second shock. It could not, I think, be maintained that the energy brought to a focus at O by these waves would in any case suffice to produce a noticeable second shock, but it may suffice to pull a trigger. If there is going to be a second shock at O, it is most likely to occur after time t_1 or t_2, or possibly some multiple of these times.

This aspect of the question has been suggested largely by Prof. Turner's claim to have discovered a period of 21·001548 minutes in the recurrence of earthquake shocks from the same epicentre.* We notice, however, that the shortest period which is suggested by our discussion is equal to the time taken for the C-waves to pass right round the earth. If we take the velocity of these waves to be 5·25 km. a second, as determined from observations of the recent Oppau explosion, we find $t_1 = 127{\cdot}2$ minutes. If we take the velocity to be 5·38 km. per second, as calculated directly from the elastic constants, we find $t_1 = 124{\cdot}1$ minutes. The average of these is 125·7 minutes, or nearly six times the period suggested by Turner. The velocity of the B-waves deduced from the earth's elastic constants is 2·99 km. a second, giving $t_2 = 223{\cdot}5$ minutes.

Our analysis suggests the possibility of two distinct periods, t_1, t_2, so that, after a shock at time $t = t_0$, subsequent shocks would be most likely to occur at the instants

$$t = t_0 + n_1 t_1 + n_2 t_2,$$

where n_1, n_2 are integers. The British Association Report for 1922 records a series of shocks in July, 1917, having the common epicentre 6·0° S., 136·0° E., which seem to fit this formula very well on taking

$$t_1 = 125{\cdot}8 \text{ mins.}, \quad t_2 = 222{\cdot}0 \text{ mins.}$$

The predictions of this formula and the times of the observed shocks are shown in the tables, Series I and II, on p. 574.

If we omit the first earthquake altogether, as having occurred almost two days before the rest, the formula predicts the time of occurrence of the remaining nine earthquakes to within an average error of 0·5 minutes for the first series and 1·6 minutes for the second series.

---

* Brit. Assoc. Report (Seismological Investigations Committee), 1922.

Series I (t_0 = 1917 July 29, 21 h. 52·4 m.).

n_1.	n_2.	$n_1t_1 + n_2t_2$.	Earthquake observed.	Error.
−22	0	−2767·6	[−2776·2]	[−8·6]
0	0	0	−0·3	−0·3
2	0	251·6	251·6	0·0
6	0	754·8	755·6	0·8
14	0	1761·2	1760·8	−0·4

Series II (t_0 = 1917 July 30, 4 h. 19·6 m.).

n_1.	n_2.	$n_1t_1 + n_2t_2$.	Earthquake observed.	Error.
0	0	0	1·4	1·4
3	0	377·4	376·6	−0·8
1	2	569·8	570·2	0·4
4	1	725·2	722·9	−2·3
6	0	754·8	756·2	1·4

A further series of twenty-one earthquakes, recorded in the same Report, fits the formula with somewhat less accuracy. These earthquakes appear to fall into two distinct series, as follows, the numbers in brackets denoting the error of the formula when we use the same values of t_1 and t_2 as before :—

I (0·0), V (4·7), VI (0·2), VIII (2·2), IX (−4·9), X (−4·3), XII (−1·0), XV (−0·5), ...

II (0·0), III (−1·1), IV (5·9), VII (1·2), XI (−2·0), XIII (3·2), XIV (4·8), ...

The average error of the formula is now 2·5 minutes in the first series and 3·0 minutes in the second.

The evidence is no doubt nothing like strong enough to establish the reality of the two suggested periodicities, but it may be worth putting forward, such as it is, in the hope that it may lead to similar examinations of further earthquake records.

The Coincidence Method for the Wave-Length Measurement of Absorption Bands.

By H. Hartridge.

(Communicated by Prof. J. N. Langley, F.R.S. Received July 3, and in revised form, November 6, 1922.)

Introduction.

It was shown by Bryan and Baker (5) that the coincidence method is one of the most accurate of visual methods of measurement. Their observations applied to sharply defined lines of various types. I have found that nearly the same degree of accuracy is obtained when the lines set into coincidence are not sharp, but have blurred edges, like those of absorption or interference bands. The reason for blurred lines being set into coincidence with nearly the same accuracy as sharp ones is to be described in detail elsewhere, with the various physiological factors on which accuracy depends.

The method of coincidences has been applied both to direct visual measurements of absorption bands and also in a few cases to photographic records of such bands. In this paper the method of direct visual observation will be considered alone, because more experience has been gained in this connection. It is possible that in future the application to photography will prove the more important because the method would seem to be applicable to such cases as stellar and mass spectra photography.

(1) The Theory of the Method.

It has long been the practice to tabulate the mean wave-lengths of the absorption bands exhibited by different pigments. It is now known, however, that even bands which are symmetrical are not also symmetrical to the eye unless the apparent edge of the band occurs at points which have identical absorption coefficients and unless the intensity of the background is uniform. In all other cases the point of greatest density will not be situated midway between the apparent edges of the absorption band. It should be noted further that there will also be a slight difference introduced if the band under examination be situated in a spectrum produced by refraction instead of one produced by diffraction.

In spite of these facts, the term mean wave-length (used to denote the wave-length midway between those of the apparent edges of the bands) has certain advantages, particularly when the bands are apparently symmetrical and the shift that they undergo under various conditions is the subject of experiment. The term mean wave-length will therefore be used in this paper in this sense.

Owing to the effects of simultaneous contrast, the appearance to the eye of an absorption band is that of a broad uniform area of absorption with no clearly marked maximum. The edges are indefinite, not corresponding usually to the place on the curve where the slope is maximum.

The edges are found to vary greatly with the concentration and thickness of the absorbing agent, and the brightness and quality of the source of light. But the effect of these factors can be eliminated during any one experiment by suitably standardizing them. Accurate measurements of the mean wave-lengths of absorption bands are found to be difficult to obtain with an ordinary spectroscope for two reasons :—

(1) If the attempt be made to set the cross wire in coincidence with the centre of the band, then the eye has to subdivide an almost uniform area of considerable breadth bounded on both sides by indefinite margins. This procedure is found to offer considerable difficulty.

(2) If, on the other hand, the measurement of the positions of the indefinite margins be attempted, then it is found that individual readings differ greatly from one another, and therefore their mean value is subject to considerable error.

(2) OPTICAL PRINCIPLES INVOLVED.

The spectroscope to be described employs a third method for measuring the mean wave-length of absorption bands. In place of the ordinary cross wire it makes use of an absorption band of the same width, mean density and distribution of absorption as that to be measured but *reversed**—a process which it is found to carry out with considerable accuracy. The absorption band provided as an index is formed in a second spectrum reversed in direction to the first, and both spectra have the same brightness and dispersion, and are both formed from light that has passed through the pigment to be examined. The bands in the index spectrum are therefore similar to those in the other, no matter what may be the concentration and thickness of the pigment, or the properties of the source of light. But the instrument is so constructed that the index spectrum may be shifted parallel to its long axis by the manipulation of the micrometer mechanism so that the two similar bands may be moved until they appear to the observer to be in line. The micrometer index then points on the scale to the mean wave-length of the band that has been the subject of measurement. This may be more clearly explained as follows: Suppose, in the first place, that there are two single bright lines in the two spectra (*e.g.*, the green Thallium lines) and that by manipulating the micrometer screw of the spectroscope these have been set into coincidence, then it is clear that, if the scale attached to the micrometer has been properly graduated, the wave-length of the Thallium lines will be indicated. Suppose in the second place, that the two bright lines are replaced by two absorption bands, and that these be set so that to a certain observer the left-hand edge of one band is in coincidence with the left-hand edge of the other band above it, then it is clear that the right-hand edges must also be in coincidence since the bands are mirror images of one another. Since the spectra are produced by diffraction, they are "normal" spectra, and, therefore, if λ' is the wave-length of one edge of the band and λ'' that of the other edge, then the point midway betwen the edge will have a wave-length of $\frac{1}{2} \times (\lambda' + \lambda'')$. Since when the edges of the bands are in coincidence the midway points must also be in coincidence, it is clear that the wave-length of $\frac{1}{2} (\lambda' + \lambda'')$ of one spectrum is in coincidence with that of the other, and, therefore, the reading given by the micrometer index will be that of $\frac{1}{2} (\lambda' + \lambda'')$.

---

* The process that the eye carries out is to judge when, *e.g.*, the long wave-length edge of one band is brought into coincidence with the short wave-length edge of the other.

(3) CONSTRUCTION OF THE INSTRUMENT.

The spectroscope is so designed that two similar spectra are provided lying side by side and reversed in direction to one another. Such an arrangement might be attained by fitting a reversing prism over half the aperture of a prism spectroscope; it is found, however, that the desired result is more efficiently obtained by the employment of a diffraction grating, the optical arrangement being shown in the figure. Light from the source L is concentrated by means of the simple lantern condenser C, on to the adjustable slit of the spectroscope S, through which some of it passes to the achromatic lens P, by which it is rendered parallel. The light now takes two routes, half of it having passed through the prism of small angle D, which deflects it slightly upwards (the object of this deflection will be considered immediately), reaches the surface of the grating G, by which a spectrum is formed which appears to the observer to lie in the direction N normal to the grating's surface, the violet end of the spectrum being to the left, and the red end to the right. The other half of the light from the collimator is deflected downwards by another prism of small angle D', and then reaches the surface of the stainless steel mirror M, by which it also is reflected on to the grating, but from the opposite side of the normal. In this manner a second spectrum is formed which appears to the observer to lie in the direction N normal to the grating's surface, but with the violet end of the spectrum to the right and the red end

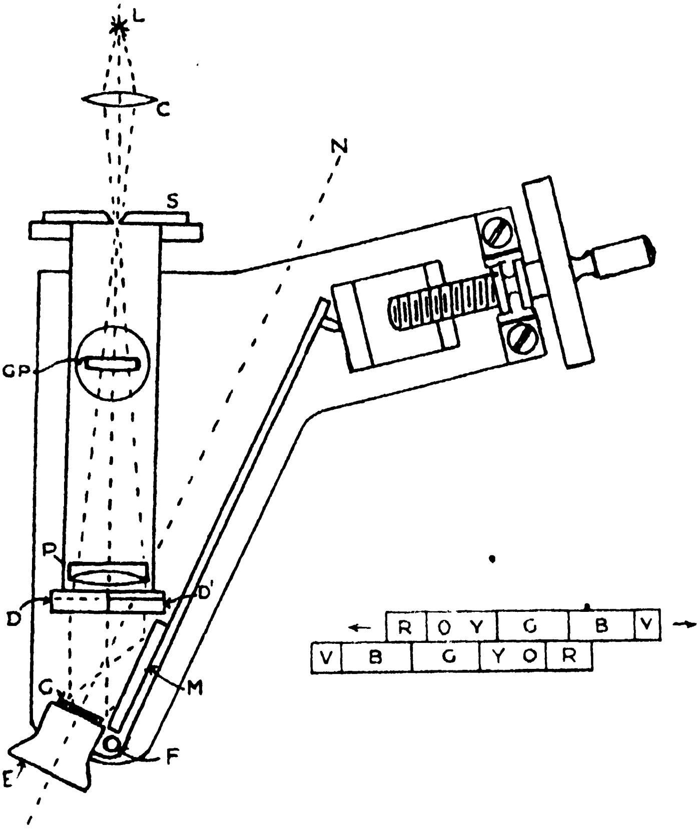


Diagram of Reversion Spectroscope.

L = Light source. C = condenser. S = adjustable slit. G.P. = glass-plate micrometer. P = achromatic lens. D & D′ = weak prisms, base up and base down respectively. G = grating. F = fulcrum. M = stainless steel mirror. E = eyepiece. N = normal to grating. The spectra are seen side by side, one reversed in direction to the other; the upper one moves horizontally on turning the micrometer screw.

to the left, *i.e.*, in the reverse direction to the spectrum previously considered. If it were not for the interposition of the weak prisms D and D′ these two spectra would lie over one another, but the prisms are given such an angle that these spectra just lie side by side (1).

The two spectra are caused to shift relatively to one another by one of two micrometer mechanisms. When a large range of wave-lengths has to be covered, the spectra are shifted by rotating the grating and mirror about the fulcrum F, by means of the micrometer screw shown in the diagram. When, however, a small range of wave-lengths has to be covered with greater accuracy, then the glass plate micrometer G is employed (2). This consists of a thin flat glass plate attached to a pointer. When the plate is set at an angle to the beam passing between the slit S and the lens P, the beam is deviated to one side, just as if the slit itself had been given a slight side shift, and this causes the two spectra to undergo a small side movement, relative to one another.

The amount of deviation produced by the plate can be calculated according to the approximate formula :-

$$D = T \tan i \left(\frac{\mu - 1}{\mu}\right).$$

Where D = deviation, T = thickness of plate, μ = refractive index and i = angle plate with incident rays.

(4) CALIBRATION OF THE MICROMETERS IN WAVE-LENGTHS.

For the measurement of absorption bands a relatively wide slit and a low degree of magnification of the spectra is usually advisable, as will be described later. But for purposes of calibration a narrow slit and a telescope of 8 or 10 diameters magnification should be used. This is preferably attached to the base of the instrument on a separate mount to the grating and mirror, lest any strain might be put on the micrometer mechanism. The latter is so designed that motion proportional to the sine of the angle is given to the frame carrying the grating and mirror. Since the wave-length varies with the sine of the angle, the graduations on the drum should correspond approximately to Angström units (1 division on the drum being equal to 1 A.U.). The graduations should be checked, however, against standard lines. A convenient source for the purpose would appear to be the new "Cathode Glow" lectric lamp, but any of the more usual sources may be used. From the readings given by the instrument for these standard lines, an "error curve" can be drawn on squared paper, which is kept for correcting future observations. The calibration of the glass plate micrometer must be effected for several different spectral regions, because its deflection is greater for light of short wave-length than for light of long. The process can be carried out either by taking a pair of standard lines (*e.g.*, the Na lines) and measuring the movement in scale divisions required to set first one pair and then the other pair into alignment. (This is a highly accurate method, and is therefore the one that should be used wherever possible.) The second method is to calibrate the glass plate micrometer from the previously calibrated screw micrometer in the following way :—

Suppose, for example, that it is required to calibrate the glass plate for the given region near the mercury line ($\lambda = 5461$), then the first operation is to set this line in both spectra in agreement. Next the glass plate micrometer is rotated as far as it will go one way and the screw micrometer again read. In general there will be a difference between this reading and the previous one, of a whole number and a fraction of drum divisions (say 3⅜). Take the whole number ($= 3$), and set the screw micrometer first 3 drum divisions on one side and then 3 on the other side of the true reading, and each time with the glass plate micrometer bring the lines back into agreement.

Since 6 divisions on the drum (corrected, if necessary, from the error curve) equal 6 A.U., the number of glass plate micrometer divisions equal to 6 A.U. is thus found. This process is repeated for any other parts of the spectrum in which it may be necessary to use the glass plate micrometer for measuring differences in wave-lengths.

The following may be quoted as an example of the latter method of calibration, carried out on a larger instrument than the one illustrated in fig. above. In this instrument 1 A.U. subtended about five minutes at the eye. 26·97 glass-plate micrometer units were found to equal 84·667 Angström units.*

The differences between the wave-lengths of the Na lines (5890 and 5896 A.U.) was found to measure 1·910 G.P. micrometer units. Therefore difference in wave-length of the two Na lines

$$= \frac{84\cdot667 \times 1\cdot91}{26\cdot97} = 5\cdot996 \text{ A.U.} \pm 0\cdot03 \text{ A.U.}\dagger$$

With a different glass-plate micrometer the values obtained were that 63·13 G.P. units are equivalent to 159·8 A.U.

The difference in wave-length of the sodium lines was found in this case to measure 2·348 G.P. units :—

$$\text{Therefore} = \text{to } \frac{159\cdot8 \times 2\cdot35}{63\cdot13} \text{ A.U.} = 5\cdot945 \text{ A.U.} \pm 0\cdot03 \text{ A.U.}$$

---

* The glass-plate micrometer units were measured by means of a tangent scale graduated in ½ mm. This scale was corrected so that the true deviation produced in the light rays by the glass plate could be measured.

† Since in this instrument 1 A.U. subtended at the eye 5 minutes of arc approximately, and since the accuracy of coincidence measurement is found experimentally to have a limiting value of 9·5 seconds, the accuracy cannot exceed 0·03 A.U. approximately.

The two values obtained for the difference in wave-length of the Na lines differ from one another by less than 1 per cent., and their mean 5·97 A.U. may be compared with the value of 5·965 A.U. obtained by Fabry and Perot by interference methods.

(5) THE QUANTITATIVE ESTIMATION OF PIGMENTS BY THE WAVE-LENGTH MEASUREMENT OF THE ABSORPTION BANDS OF THE SOLUTION CONTAINING THEM.

Pigments are generally estimated quantitatively by one of two methods: (1) By an estimation of their depth of tint by means of a colorimeter; (2) by an estimation of their absorption relation in some chosen parts of the visible, photographic or radiant heat spectrum.

It has been found that under certain circumstances a third method exists, which may give results of considerable accuracy. This method depends on the principle that if the absorption bands of two pigments have different mean wave-lengths then if both pigments are in the same solution and are together causing absorption of light, the measurement of the mean wave-length of the absorption bands will be found to vary with the relative concentrations of the pigments if (1) the band of one pigment is approximately similar in the form of its density curve* to a band of the other pigment, and if (2) at no concentration do the bands appear separate; that is, that their mean wave-lengths must not differ by more than the apparent mean widths of the bands.

In the application of the method either one or both pigments may vary in concentration. The pigments may be in the same glass vessel or in separate glass vessels through which the light passes in turn, or the pigments may be in dyed collodion or gelatine films. For example, suppose it is desired to estimate the concentration in solution of a pigment A. A pigment B is selected so that its absorption bands and those of A are related to one another in the manner specified above.

Various known mixtures of these pigments are prepared; they are submitted to spectroscopic examination and the "mean" wave-length of the most suitable absorption band is plotted against the relative concentrations of the pigments. Alternatively a double wedge trough is used, into each side of which one of the pigments is placed. When a beam of light is passed through this trough, and the light subjected to spectroscopic examination, the effect of different relative concentrations of the pigments may be obtained.

Having obtained a calibration curve by one of these methods, an unknown amount of either pigment may be estimated by adding to the solution a known concentration of the other pigment, and then applying the reversion spectroscope for measuring the mean wave-length of the resulting absorption bands.

---

* Absorption coefficients plotted against λ.

It should be pointed out that, for the purpose of determining the relative concentration of two pigments, a purely empirical micrometer scale would suffice.

Since certain relative concentrations of two pigments require certain settings of the micrometer, any subsequent setting required by a solution under examination could be referred to the previously ascertained relationship between relative concentration and setting for the relative concentration to be ascertained. An example of the method will now be given, because other estimations could probably be based on a similar technique.

(6) AN EXAMPLE OF THE USE OF THE METHOD.

The Estimation of the Percentage Saturation of Blood with Carbon Monoxide by the wave-length measurements of its absorption bands (1).

It has long been known that the two absorption bands of blood, in the visible spectrum, are shifted towards the violet if carbon monoxide gas be substituted for oxygen.* The ordinary spectroscope demonstrates the shift well, but is insufficiently accurate for a method of estimation to be based on its measurements. The greater accuracy obtained by the reversion spectroscope, enables the relative amounts of the two pigments, oxyhæmoglobin and carbon monoxide hæmoglobin, to be determined to within about 1 per cent. Since the absorption bands shift approximately 60 A.U., the wave-lengths of the bands are being determined with an error of about 0·6 Angström units. It should be pointed out in the first place that the bands appear to the eye to shift, because while the "density curves" of the bands overlap one another their summits occur at different wave-lengths. At any given percentage saturation, except 0 and 100, both pigments are present together in solution, and, therefore, both absorb the light as it passes through them. The density curve of the mixture will therefore be the summation of the density curves of each pigment separately, the relative heights of the summits of the two curves above the base line being in proportion to the amounts of the two pigments present. If such a summation be effected graphically for different proportions of the pigments, it will be found that the mean wave-length of the summation curve varies in correspondence. Of the two bands in the visible spectrum given by blood pigment, the one near the D lines is sharper and narrower, and is therefore the one that is measured when estimating the carbon monoxide saturation. To ascertain the wave-length of this band for different percentage saturations of blood with CO, known mixtures of solutions of oxy and carbon monoxide hæmoglobin might be prepared for spectroscopic analysis, if it were not for the fact that the water in which the blood pigment is dissolved itself takes up an unknown amount of the gas. The mixtures may therefore actually have a very different proportion of the two pigments from that calculated from the volumes of each taken. The fact is, therefore, made use of, that the absorption bands will be the same whether the two pigments are in two separate vessels through which the light passes in turn, or whether they are mixed together in the same vessel. A rectangular glass trough is divided by a diagonal partition into two separate wedge-shaped troughs. Into one is placed some blood pigment saturated with oxygen, and into the other some of the same solution fully saturated with CO gas (a few drops of ammonium sulphide being added in order to remove the free oxygen from the solution). A beam of light in passing through different parts of the double trough encounters the same total number of molecules of blood pigment, but a different relative thickness of those combined with oxygen to those combined with CO. From the thicknesses of the two solutions the effective percentage saturation with CO can be at once calculated. This trough, filled as described, is placed before the spectroscope, and the wave-lengths corresponding to the different percentage saturations with carbon monoxide are determined, and plotted on a graph. This graph is used for determining the percentage saturation of different blood samples with CO gas, of which the wave-length of the α-band has been estimated with the reversion spectroscope. The form of the graph has been previously given elsewhere (1).

---

* This shift is due to the fact that the combination between the blood pigment and oxygen, viz., oxyhæmoglobin, has absorption bands nearer the red than those of the carbon monoxide compound.

The following experiment provides a searching test of the degree of accuracy which the method is capable of yielding. The α-band of oxyhæmoglobin being made the object of measurement, one observer (H. H.) set the micrometer screw of the spectroscope to different wave-lengths in turn which were unknown to the other observer. The latter (F. J. R.) reported on the apparent closeness of the coincidence of the bands. The results obtained are shown in the following table :—

Wave-length	5764	5765	5766	5767	5768	5769	5770	5771
Apparent closeness		Not	Close	Very close	Close	Not	Not	
				Not	Not quite	Very close	Not quite	Not
	Not	Not	Not	Not quite	Very close	Not quite		

It will be seen that 5765 and 5771 were definitely excluded as being too low and too high respectively, leaving a possible range of 5 A.U. Of this range, the middle value is clearly the correct one from the answers given. If all wave-length settings to which an answer of "close" or "not quite" are included, the inaccuracy is not even then found to be greater than 2 A.U. When the observer at the spectroscope himself sets the bands into coincidence by means of the micrometer screw the difference in wave-length between two separate sets of ten readings seldom differs by more than 1 A.U. Thus the mean wave-length of the *a*-band of the solution used in the above experiment was measured by the same observer using the usual method, and was found to equal 5769 A.U. To further test the method, duplicate determinations on twelve samples of sheep's blood were made using (*a*) the spectroscope and (*b*) the blood gas pump. The values given in the following table were obtained :—

Blood gas pump	10·1	17	25·3	35	40·1	40·7	45·4	47·2	68·5	69·8	78·7	82·5
Spectroscope ..	12·0	17	22·0	34	39·0	39·0	42·0	48·0	74·0	74·0	71·0	80·0
Difference ..	+1·9	±0	−3·3	−1	−1·1	−1·7	−3·4	+·8	+5·5	+4·2	−7·7	−2·5

It will be seen that there is fair agreement between the values obtained by the two methods, particularly when it is remembered that the possible sources of error in the case of the blood gas pump are considerable. For values of the percentage saturation of hæmoglobin with CO of under 60 per cent., the average difference between the two methods was 1·6 per cent.

(7) ACCURACY AND SOURCES OF ERROR.

Measurement has shown that absorption bands vary in their mean wave-lengths under the following conditions :—

(1) Change of temperature(4). In the case of the α-bands of the oxy and carbon monoxide compounds of blood, the movement is 0·28 A.U. per °C. towards the red.

(2) Change of light source(3). A relative increase in the light intensity on one side of an absorption band is found to shift the band towards the darker side.

(3) Change in the observer. Experiment shows that the readings of one observer are quite consistent with one another, while the readings of two different observers may differ considerably.

It is necessary, therefore, for an observer to obtain a calibration curve for the spectroscope for himself, and to give if possible always his measurements in reference to a line or band of known mean wave-length, the temperature being recorded at which the observations are made.

If these precautions are taken, it is found that absorption bands which have a mean width of about 120 A.U. are determined if a mean of 10 readings be taken with an accuracy usually greater than 1 A.U., and that the individual readings do not usually diverge from the mean value by more than 2 A.U. As the limit to accuracy depends on the eye, it might be well to compare the above values with those obtained by other methods of measurement.

The coincidence of two black lines was measured by Baker and Bryan(5), and found to have an average error of 9·5 seconds of arc. The bisection of the space between two black lines by a third line was measured by the same observers and found to have an average error of 8·1 seconds of arc.

In the case of the reversion spectroscope with a grating with 14,000 lines per inch a difference in wave-length of 1 A.U. is found to subtend at the eye an angle of 10 seconds of arc. The average error in the coincidence of two absorption bands of considerable width (1200 seconds of arc), and with indefinite edges would therefore appear to be of nearly the same order as that obtained in the case of two sharp lines, a conclusion of considerable interest from the physiological point of view.

It might be thought that the accuracy of absorption band measurement would be increased by increasing the magnification ; such, however, was not found to be the case, presumably because of the greater indefiniteness to the eye of the apparent boundaries of the absorption bands that are being measured.

I should like to express my thanks to Mr. A. E. Milne for his assistance in the revision of this paper.

Summary.

(1) Measurements of the absorption bands of pigments by means of the ordinary spectroscope are found to be inaccurate, because of the breadth of the bands and the indefiniteness of their margins.

(2) The adjustment of two similar absorption bands into coincidence is found to be effected with considerable accuracy. If then a spectroscope is designed in which two spectra are seen side by side, on looking down the eyepiece, but reversed in direction with one another, the measurement of the mean wave-length of the absorption bands can be accurately carried out.

(3) The optical construction of the instrument is described in Section 3.

(4) The method of calibrating the micrometers of the spectroscope are described and experimental values are given to show the accuracy obtained.

(5) The method for the quantitative estimation of pigments by the wave-length measurement of their absorption bands is described. The estimation is shown to depend on the movement of the bands which occurs when the concentration of one pigment changes.

(6) An example of the method is given, namely, the estimation of the percentage saturation of blood with CO gas by the measurement of the wave-length of the α-absorption band. The accuracy of measurement is found to be approximately 0·6 A.U.

(7) The sources of error of the method are considered, and the probable accuracy of measurement discussed. The probable error of setting two absorption bands into coincidence is found to be but little greater than that of setting two sharp black lines into coincidence or of making one line bisect the area between two other lines.

[*Note.—Added November* 25, 1922.—The referee to whom this paper was submitted has pointed out that the " mean wave-length " found by the method described in the paper above will in general differ slightly from the true maximum of absorption.

The actual error, however, will depend in the apparent width of the band. He suggests, further, that if the apparent width of the band were varied, *e.g.*, by varying the concentration or thickness of the absorbing substance, it might be possible to find a simple relation between the mean wave-length of a particular band and its apparent width, and by extrapolating to zero thick-

ness, or concentration of the absorbing substance, the true maximum might be found with considerable accuracy.]

REFERENCES.

(1) 'Journ. Physiol,' p. 1 (1912).
(2) Poynting in 'Optical Convention,' p. 79 (1905).
(3) 'Roy. Soc. Proc.,' B, vol. 86, p. 128 (1912).
(4) 'Proc. Physiol Soc.,' Feb. (1921).
(5) Bryan and Baker in 'Optical Convention,' p. 264 (1912).

On the Changes in Iron and Steel at Temperatures below 280° *C.*

By FRANK CHARLES THOMPSON, Professor of Metallurgy, University of Manchester, and EDWIN WHITEHEAD.

(Communicated by Prof. H. C. H. Carpenter, F.R.S. Received August 15, 1922.)

As a result of some unexpected facts observed during an investigation on the etching properties of the carbides in alloy steels, the study of the changes which are well known to occur in iron carbide was commenced. Since pure iron carbide was not to be obtained in bulk, the material used was a Swedish white pig iron which contained about 60 per cent. Fe_3C. The remainder consisted of more or less pure iron, and in view of the complexity of the curves expressing the relationship of the electrical resistance to temperature it was necessary to add to the high carbon material as pure a sample of iron as could be obtained. This was a remarkably pure sample of American ingot iron containing 99·9 per cent. iron. To link up these two materials, two steels with about 0·15 and 1 per cent. of carbon were also examined.

The full analyses of the materials used are recorded in Table I.

Table I.

Material.	C per cent.	Si per cent.	Mn per cent.	S per cent.	P per cent.
Iron	0·02	0·01	0·02	0·023	0·008
Mild steel	0·19	0·13	0·38	0·067	0·053
High carbon steel	0·99	0·47	0·39	0·069	0·09
White pig iron	3·8	0·2	0·35	0·02	0·02

It is now well known that carbide of iron, Fe_3C, has a change point at about 200° C., but the fact that there are also equally important changes in pure iron

has not, hitherto, been clearly realised, although there have, from time to time, been found values for various physical and mechanical properties which did not fit well on curves showing the general relationship of the property and temperature.

The first discovery of any definite change in carbide of iron appears to have been made by S. W. J. Smith in 1911 (1), during work on the influence of temperature on the magnetic properties of carbon steels. As early as 1856, Fairbairn (2) realised the importance of investigating the mechanical properties of iron and steel at temperatures between the normal and a dull red heat, in connection with the permanent setting of boiler plates. His results are now of mainly historical interest, since the methods of measuring temperature were not very reliable. In 1890 Martens (3) published the results of an important enquiry into the properties of mild steel up to a temperature of 600° C.

Since then the mechanical properties at temperatures above the normal have received repeated attention, notably by Dupuy (4), Robin (5), Huntingdon (6), and Lea (7). Many other physical properties have also been studied, the most important case from the point of view of the present work being that of the acoustical properties by Robin (8). Much work has been carried out by Honda (9) and his co-workers on the change of cementite at 200° C., which, however, suffers from the fact that they have not realised that iron itself is also changing in the same temperature range.

The results of these investigations so far as they concern the present research will be discussed when the authors' conclusions have been given.

This work consists of measurements of the electrical resistivity and thermo-electric power of the materials of Table I at temperatures from 20 to 280° C. Although some indications of thermal evolutions have been obtained, no attempt to study these has been made, as a result of the fact that work on these lines has been done at the University of Sheffield by Mr. F. Adcock, under the supervision of Prof. C. H. Desch. Several small heat changes have been observed in pure iron which correspond fairly well with the abnormalities in the electrical resistance obtained by the present authors.

Electrical Resistivity.

The measurements were made in one series by merely observing the deflection produced by passing a constant current through the wire or bar of the alloy, and are, therefore, merely qualitative, and in the second series by measurement with a Kelvin double-bridge instrument manufactured by Paul. Since it was the changes in the resistance with temperature, rather than the actual values, which were desired, it has not been considered necessary to convert any of the readings to absolute values.

Before any measurements were made the materials were reheated to 300° C., a temperature above the highest to be subsequently used, and the first heating and cooling curves have been neglected as being unsatisfactory. In addition, all the materials, with the exception of the Swedish white iron, which was in the cast condition, were annealed, before their receipt, at 650° C., which entirely removed any cold work to which they had been subjected, and left them in a soft state, showing no traces of any distortion of the normal structure.

The iron or steel was heated in an oil bath heated by gas. The oil was kept well stirred, and the temperature measured by a mercury thermometer. No correction had been applied for the exposed stem. The leads carrying the current were soldered to the wire or bar.

Since many small alterations in the rates of heating and cooling were observed at various temperatures it was deemed advisable to check the arrangement against a material which was almost certainly free from any abnormality in the temperature range under consideration. Pure copper was used for the purpose, which had been fully annealed, and showed nothing but equiaxed twinned crystals; and although in the first heating the curve obtained was somewhat irregular, on repeated heating this disappeared, and a smooth curve, free from any sudden alterations of direction, was obtained. The fact is of some interest from the point of view of the supposed allotropy of copper, which, at any rate so far as the electrical resistance is concerned, is not evidenced. The reason for the slight abnormalities at first shown is probably to be found in the small amount of cold work put upon the metal in the process of coiling it in the oil bath.

For the sake of greater clearness, the results have been plotted in the form of the change of resistance for 1° C. rise or fall of temperature against temperature, and correspond more or less with the "Inverse-Rate" method used in thermal work.

The results for the four materials used are shown graphically in fig. 1. The curves are so remarkably irregular that, were it not for the fact that they could be reproduced time after time, the points would have been put down to experimental error. On some of the materials no less than fifteen complete curves were obtained, and the points shown were obtained time after time. Further, it will be shown that there are clearly marked gradations as the carbon content of the alloy is raised, that other properties show similar results, whilst finally the work has since been repeated using entirely different experimental conditions.

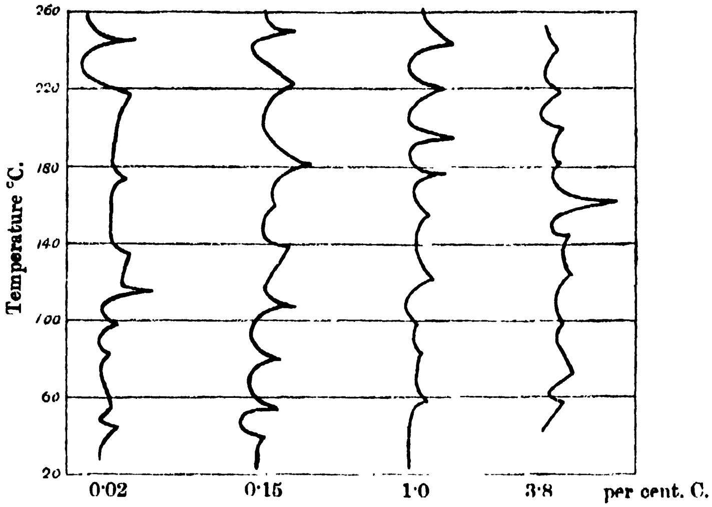


Fig. 1.

From these curves one may conclude that in the iron and in the low carbon steel there is a change proceeding almost from room temperatures. This gives a peak at 43° C. Although, as will be seen, almost all of the iron points are still to be found in the higher carbon alloys, this point is absent. It is still uncertain how far the iron which forms one of the constituents of the eutectoid—pearlite—is identical with structurally free ferrite. There are, however, some points which suggest a difference between them, and this absence of the peak at 43° C., which is quite strongly marked in iron lends weight to the view that there is some solubility of carbide in the iron of the pearlite complex. This is almost a confirmation of the view held by Benedicks of the occurrence of the modified form of ferrite, to which he has given the name " ferronite."

All the curves show a point at 55° C., which is, there can be little doubt, due to iron. There is some ground for the belief that this change in iron exerts a considerable effect on its mechanical properties. At 82° C. there is a further indication of change in the iron. It is still evident in the case of the white iron though at a slightly lower temperature and spread over a larger range of temperature, a further confirmation of the fact that in pearlite the iron has a much greater solubility for carbon or other impurities than has iron which is structurally free. In most of the curves there is at, or around, 100° C. a small point.

The peak around 120° C. shown in all of the curves is from the point of view of its importance in connection with the mechanical properties of the alloys, one of great importance. There is abundant reason for the belief that at this temperature iron undergoes a fairly profound change. This will be referred to again in connection with Robin's work on the acoustical properties of the materials.

The low carbon steel and the iron show a peak, which may also be present at a rather higher temperature in the white iron. It would be expected that as the purity of the material decreased the temperature at which the points occurred would be lowered. There is no indication that this is the case, in fact the very reverse appears to be true. This may be of some importance in connection with the explanation of these abnormalities, since the fact is opposed to an explanation in terms of allotropy.

The first clear indication of the influence of the carbide is shown in the point at 160° C. The pure iron shows no sign at all of this change. In the mild steel there is a small but quite distinct peak, while as the carbon is still further increased the size of this progressively rises until with 3·8 per cent. carbon the peak is the largest shown by the alloy. This point is perhaps one of the best proofs available that the curves do really represent definite alterations in the materials. Attempts to simplify the curves by evening up the points observed have been unsuccessful, and to the authors there is a settled conviction that even so apparently simple a metal as iron is capable of revealing surprising complications. An iron point makes itself evident at around 170–180° C. in all the materials used.

At 200° C. the well-known carbide change reveals itself in the curves for the 1·0 and 3·8 per cent. alloys. It is absent from those of the iron and the mild steel. It is of distinct interest that in the white iron the peak at 200° C. is much smaller than that at 160° C.

A further very important iron change takes place at 220° C. It is clearly shown in all the curves, though the size naturally decreases as the carbon content is raised. It is of real importance in connection with the mechanical properties of the steels and iron, and further work at present in progress has revealed an important influence on the elastic-limit of the material.

The last point observed in the temperature range investigated is at 240–250° C., and is due to the iron. It is fairly certain that had the work been carried up to still higher temperatures, further abnormalities would have been detected. Since the prime reason for the present investigation was the study of the change in cementite there was, for the present, no justification for extending the range of temperature beyond 280° C.

The results of this part of the work may be summarised as follows :—

Iron shows several abrupt changes in the rate of change of electrical resistance with temperature, of which the more important are at 55, 120, 220 and 245° C. The numerous smaller points may be either extensions, or doublings of these, or possibly to some extent the result of slight cold work produced by bending the wires, as was shown to occur in the case of copper.

Two well-marked points occur in the high carbon alloys, due to the change in the carbide of iron, at 160 and 200° C. The former is at least as important as the latter from the point of view of its effect on the resistance. How far these are distinct, or the commencement and end of a single transformation, it is not at the present time possible to say. It is worth while pointing out that since iron itself is undergoing changes in the same temperature range as that which includes the transformation of cementite, effects may in the past have been ascribed to the latter which are in reality due to the iron. In particular, the suggestion that the carbide change is continued up to 240° C. is not confirmed by the present work, but could be readily explained on the assumption that the iron change at 220 or 240° C. was the cause of the continuance of the apparent effects due to carbide.

All the curves shown in fig. 1 were obtained with a constant rate of heating of about 2° C. per minute. A considerable number of curves were obtained with different rates of heating and cooling. For iron by quenching an almost complete suppression of the points may be obtained, though in general iron wire undergoes the transformations with considerable rapidity, and quite rapid heating or cooling is required to effect much alteration of the general type of curve. The rate of heating of 2° C. per minute is quite sufficient to obtain equilibrium, as is evidenced by the results in fig. 2. This is the upper part of a

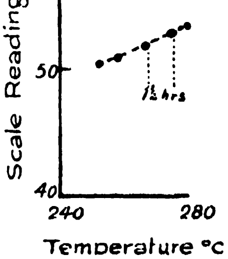


FIG. 2.

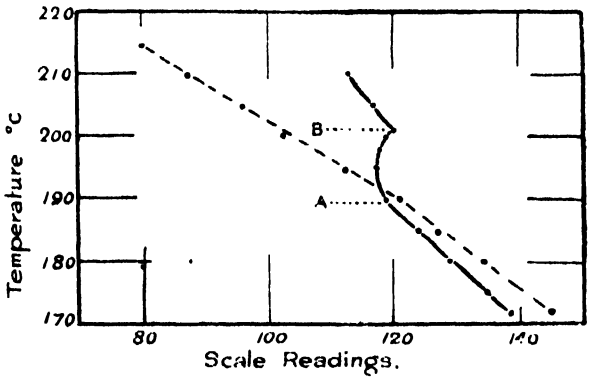


FIG. 3.

curve for the iron. The metal was maintained at a temperature of 275° C. for 1½ hours. After this the heating was continued, and the values for the resistance still fall on exactly the same line as before.

If, however, the iron is quenched, the suppression of the changes is revealed by the alteration of the resistance compared with that for the slowly cooled sample. The following figures show this effect:—

Resistance of an iron wire slowly cooled from 270° C.	..	0·3712 ohms.
,, ,, ,, quenched from 270° C.	..	0·3677 ,,

In the case of the carbide transformation, however, the influence of the rate of temperature is very much more important. The changes are far more sluggish, and are more easily suppressed. In the case of a bar of white iron about 1 sq. cm. in cross-section, even air-cooling is sufficient almost entirely to remove any indication of change. This is well shown in fig. 4. The upper curve is that of the bar cooled from 255 to 140° C. in 2½ hours, and shows the normal points. When cooled through this temperature interval in 20 minutes, however, the curve is almost smooth, a slight change of direction at around 200° C. being the only remnant left of the transformations. The effect of quenching on the resistance will be considered later.

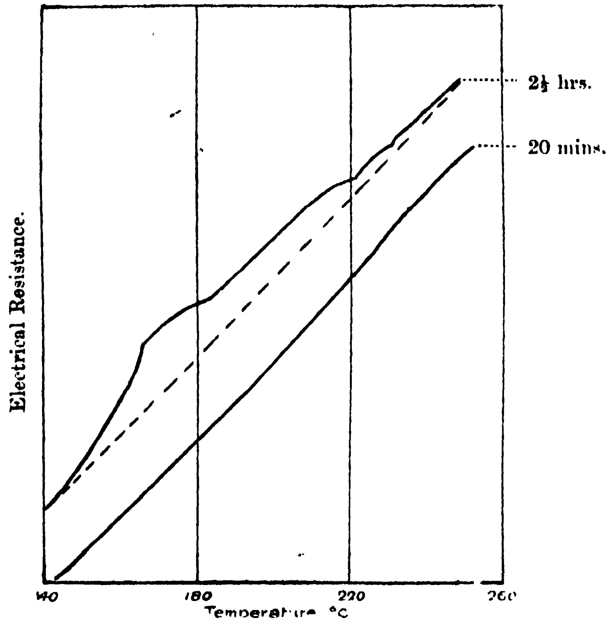


Fig. 4.

The rate of heating is of equal importance. In fig. 3 are recorded curves for the white iron heated at different rates. The dotted line is for the bar heated at a rate of 2° C. per minute from 180 to 210° C. plotted as scale readings against temperature. The full curve was for a heating in which the sample was maintained at 195–205° C. for half an hour. This extra time has permitted the carbide to transform much more completely, producing a marked increase in the resistance as compared with the more rapidly heated sample. As a result of this the slope of the curve after the heating is distinctly different from that of the more rapidly heated bar. The slowness with which the cementite is transformed is such that it appears to be a simple matter to preserve at room temperatures the form stable only above 200° C. As will be shown this is only true immediately after the quenching, since on standing it reverts to the stable condition.

This possibility of retaining the high temperature form of cementite at room temperatures by quenching received confirmation during the experiments on the etching properties of the carbide. To substantiate the conclusions drawn from some of these, two wires, one of the 1 per cent. carbon steel and the other of iron, were quenched from 280° C.

The electrical resistances were at once measured, and then subsequently for about 12 days. The results are shown in fig. 5. In each case the resistance in the slowly cooled state corresponded to that after about 12 days. It will thus be seen that the quenched wires possessed distinctly different resistivities from those which had been slowly cooled. This of course would be expected, since there are many changes which the quenching has suppressed. On standing at room temperatures the wires slowly reverted to the normal state. This

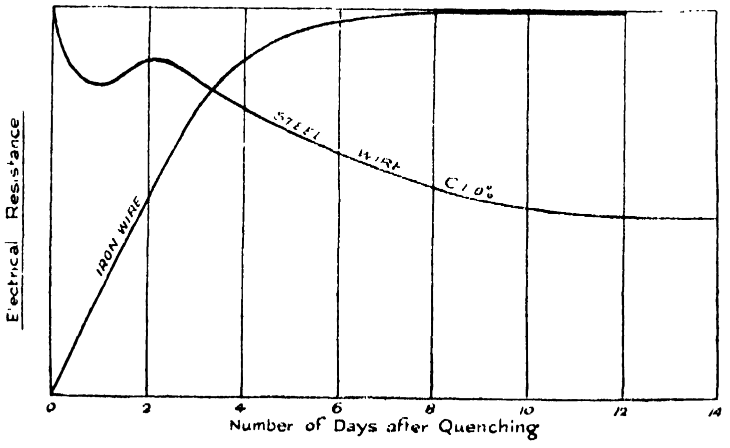


Fig. 5.

automatic tempering is almost as unexpected as the effect of quenching from these comparatively low temperatures itself.

The actual values of the resistance of the iron wire during this change were :—

Time after quenching in days :—	Resistance at 17° C.
—	0·3677 ohms.
1	0·3687 ,,
2	0·3697 ,,
4	0·3704 ,,
7	0·3708 ,,
10	0·3710 ,,
Slowly cooled, 280° C.	0·3712 ,,

The changes in the high carbon steel were :—

Time after quenching in days :—	Resistance at 15° C.
—	0·1460 ohms.
1	0·1452 ,,
2	0·1455 ,,
3	0·1454 ,,
4	0·1450 ,,
5	0·1446 ,,
6	0·1448 ,,
7	0·1443 ,,
8	0·1440 ,,
12	0·1438 ,,
14	0·1437 ,,
Slowly cooled from 280° C.	0·1439 ,,

The values for the iron are of interest in showing that the changes are not merely the effects of stresses set up by a sudden cooling apart from unusual factors. Had it been simply a question of the former which might be expected in any material, one would have been led to expect that the results of the treatment would have been to increase the resistance. It has been shown that for the iron the value for the latter is lower than it is when slowly cooled. In comparison with the curve for the iron, that for the carbon steel is of interest in that the effect of the tempering at room temperatures is in the opposite direction, the resistance decreasing as the alloy undergoes the changes. The dip in the earlier part of the curve is of considerable interest. It appears that it may be best explained on the assumption that the effects of both the carbide and the iron changes are superposed, and that for the carbide itself there would be a continuous decrease of resistance with time till the constant value had been reached. Thus the resistivity of β or high temperature phase of cementite would appear to be greater than that of α or normal carbide.

Further it will be noticed that the iron is more rapidly reconverted into the stable state than is the carbide. The greater hardness of the latter would lead one to expect that this should be the case.

The fact that the high temperature form of cementite is self-tempering at room temperatures was confirmed by some etching properties of the material. A method has been discovered by which the two forms may be differentiated ; and whereas iron carbide immediately after quenching from 280° C. reacts as the β form, in the course of a day of two the differentiation becomes much less clear, and after the sample has remained at room temperatures for a week the carbide possesses all the etching characteristics of the α state.

Since there is good reason to think that the carbide change, at least, is accompanied by a volume change, this gradual alteration of rapidly cooled samples of iron and steel, by setting up stresses may well afford an explanation of the mysterious cracks which at times form in castings on standing, especially where the material is somewhat warm. Another case in which these results may throw light on industrial failures is that of quenched high carbon steels. These are liable to crack some time after the quenching has been carried out. Though this is in part undoubtedly due to alterations in the solid solution, the residual austenite being slowly converted into martensite, there is also the possibility that the suppressed changes in the iron and the carbide may exert an influence. Immediately after the quenching any cementite must be in the β form. On standing this will revert to the α state and in so doing will undergo a change of volume. This superposed on that due to the austenite-martensite effect may set up sufficient internal stress to produce fracture.

There are two types of heat-treatment to which such high carbon tool steels may be subjected. In the more common one the quenching is effected from a temperature as little above the AcI point as possible. This will leave some undissolved cementite. In the other type the quenching is done from a temperature at which all the cementite has just passed into solution. There is much to be said for each treatment, but it is well to point out that from the point of view of the stresses set up by the reversion of β to α cementite, the latter treatment appears to possess certain advantages.

Thermo-electric Power.

The very curious results obtained in the course of the resistivity work were confirmed by a few experiments on the thermo-electric power of the iron and the high carbon steel wires.

These were made into thermo-couples against platinum, a metal which, so far as is known, suffers no change in the temperature range investigated. The couple was heated slowly in the oil-bath previously used, and the deflection of a high resistance mirror galvanometer taken as the measure of the E.M.F. produced. This was plotted as the differential. Since it was the change in this property produced by alterations in the states of the wires which it was desired to detect, no attempt has been made to convert the results into any absolute form.

In the case of the low carbon steel (fig. 6) the results are a very satisfactory confirmation of the results of the resistance work. Distinct peaks occur at 56, 98, 120, 140 and 220° C. The points at 120 and 220° C. are the largest, in agreement with the view already expressed that they are the most important.

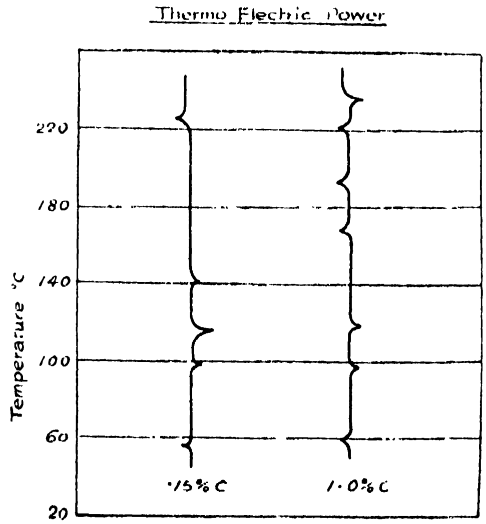


FIG. 6.

The curve for the higher carbon steel is equally satisfactory confirmation of the earlier work. As in the case of the resistance determinations, the iron points are well shown, while in addition the points due to carbide at 160 and 200° C. are clearly defined. The iron point at 236° C. is shown in this latter curve, though it is absent from the former, due possibly to the heating being discontinued a little too soon.

Broadly the results of the work on the thermo-electric power are completely confirmatory of the conclusions already drawn.

Work of Previous Investigators.

The chief results of earlier workers, so far as they affect the problem attacked by the present authors, on the changes in iron and steel at these low temperatures, are:—

The work of Robin on the acoustical properties of iron and steel has considerable interest as affording striking confirmation of the existence of a change at 120° C. in iron and its persistence in the steels. His curves are also of much value as showing the commencement of the carbide transformation at 160° C. After the low value for the duration of sound at 120° C. the curve rises, until as the next, or carbide, change sets in it again tends to fall. Further, several curves show changes in direction around 240–250° C.

From his work on the loss of ferro-magnetism in steel on heating, S. W. J. Smith has indicated that this effect due to carbide sets in at about 160–170° C. It extends over a long range, being complete at about 240° C. This temperature is much higher than that found by the present authors, and it appears to be very probable that changes in iron, which, as has been shown, occur at 220 and 240° C., are responsible for the apparent extension of the carbide transformation to this temperature.

Honda's work gives a range of about 160–215° C. for the carbide changes. His work is of much interest, in that he has determined the values for several physical properties of α and β cementite.

The mechanical properties of iron and steel at temperatures above the normal also bear witness to the fact that there are factors at work which are unusual. The influence of the two important points in iron at 120 and 220 ° C., is very well illustrated in the curve obtained by Lea for the modulus of rigidity of a mild steel. The curves of Dupuy for the tensile strength and the reduction of area of iron and a series of steels also point to the fact that this temperature range is one in which abnormal effects are present.

Thermal evolutions on cooling have been detected in both iron and carbon steels around 120, 140, 160 and 200° C. by Movius and Scott (10).

Finally the authors are informed that there are volume changes produced by some at least of the points discovered, a fact which is at present receiving further consideration.

Summary of Conclusions.

(1) Iron shows abnormalities of the rates of increase of electrical resistance and electric potential against platinum at well-marked temperatures. Below 280° C. these temperatures are: 55, 100, 120, 140, 220 and 245° C. Of these those at 120 and 220° C. are the most important. It is difficult to believe that these can be of the nature of allotropic changes, and for the time the explanation must be taken as unknown.

(2) Under the same conditions carbide of iron possess two well-marked points at 160 and 200 ° C. It is not clear whether these are distinct points or are the ends of a single transformation range.

(3) When samples of iron and high carbon steel are quenched from 280° C., it is found that the electrical resistivities differ from those obtained by slow cooling. These values are not constant, but as the material tempers they gradually alter till after the lapse of some days the values practically coincide with those obtained by slow cooling.

REFERENCES.

(1) Smith 'Proc. Phys. Soc.,' 63, 342, 1911, and 77, 1912.
(2) Fairbairn 'British Assoc. Rep.,' 405, 1856.
(3) Martens 'Abstract Proc. Inst. Civil Eng.,' vol. 104, 209.
(4) Dupuy 'Journ. Iron and Steel Inst.,' 91, 1921.
(5) Robin 'Journ. Iron and Steel Inst.,' Carnegie, vol. 2, 1910.
(6) Huntingdon 'Journ. Inst. Metals,' vol. 2, 126, 1912.
(7) Lea 'Engineering,' June 30, 1922.
(8) Robin 'Journ. Iron and Steel Inst.,' Carnegie, vol. 3, 1911.
(9) Honda and Takagi 'Sci. Rep. Tohoku Univ.,' 203, 1913, and 161, 1919.
(10) Movius and Scott 'Chem. and Met. Eng.,' 1069, 1920.

Smokes: Part I.—*A Study of their Behaviour and a Method of Determining the Number of Particles they Contain.*

By R. Whytlaw-Gray, J. B. Speakman, and J. H. P. Campbell.

(Communicated by Prof. A. Smithells, F.R.S. Received October 25, 1922.)

[Plate 5.]

Theoretically, at least, systems of solid particles of ultra-microscopic size suspended in a gaseous medium, should show many resemblances to those analogous systems in which the dispersion medium is a liquid. It might be expected that the well-known characteristics of colloidal matter, such as coagulation, peptisation, protective action, gel formation, etc., would have their counterpart in the simpler gaseous systems. Should this prove to be the case much information should accrue from a study of these aerosols* in which many complicating factors are absent.

Now although the movement of individual particles (not necessarily of ultra-microscopic size), suspended in gases, has been investigated very carefully, and the study has led to experimental and theoretical results of the greatest importance, for instance, the measurement of the unit electronic charge (de Broglie (1), Ehrenhaft (2), Millikan (3)), little attention has been paid to the question of the stability of gaseous systems and the changes they undergo, and no attempt has been made to compare the properties of solid colloidal suspensions in the gaseous and liquid states. From this point of view, dusts, fumes, smokes, and clouds are colloidal systems possessing varying degrees of stability, dispersion, and concentration, which sometimes flocculate or precipitate rapidly, sometimes remain highly disperse for long periods of time, and correspond to the many types of liquid systems which vary from coarse suspensions to fine-textured sols.

Before any correlation of the properties of the two classes of dispersoids can be attempted it is necessary to show :—

(1) That corresponding systems in the gaseous phase can be prepared and can exist for measurable periods in a stable state.

(2) To work out a reliable method of determining the number of particles in a given volume of the dispersoid.

(3) To be able to determine the average size of a particle in any given system.

---

* Aerosol is a convenient term to denote a system of particles of ultra-microscopic size dispersed in a gas, suggested to us by Prof. Donnan.

In the work about to be described we have investigated two of these three points, and we believe the method to be of general applicability for the study of these dispersoids.

Experimental.—The Preparation of the Dispersoid.

One essential point of difference between dispersoids in liquid and in gaseous media is the much greater amplitude of Brownian movement of the particles in the latter. Hence aerosols could only be expected, other conditions being similar, to show the same stability at much greater dilutions. Further, unlike solid liquid systems in which the particles all carry electric charges of the same sign, the systems with which we are dealing contain, as a rule, in addition to the electrically neutral, particles of opposite electric sign (4) which makes for instability.

Again experience has shown that highly disperse and homogeneous smokes are very difficult to prepare when the concentration exceeds a certain limit. We have therefore confined our attention to very dilute systems, in which the concentration is of the order of 20 mgrm. per cubic metre.

There are a number of ways of dispersing solids in the form of fumes or smoke (see Brit. Assoc., 4th Colloid Report, 1922 (5), and Tolman, J. (6)).

We have used three main methods :—

(1) Sudden cooling of saturated vapour by admixture with a large volume of air or gas. Almost any substance volatile up to a red heat can be treated in this way. To obtain a high degree of dispersion the substance is spread in thin layers over a surface which can be heated electrically, and volatilisation carried out in a rapid stream of air.

Such bodies as anthracene, acetanilide, diphenylamine, ortho-phthalic acid, paraffin wax, etc., yield readily by this treatment fine-textured smokes ; others such as benzyl benzoate, or clean mercury, give either a coarse dispersoid or no detectable cloud.

(2) The chemical interaction of gaseous substances to form a solid. Ammonium chloride smokes can be prepared over a wide range of dispersion by mixing the highly diluted constituent gases. In a similar way zinc oxide can be obtained in the form of a fine dispersoid by the decomposition of zinc ethyl vapour highly diluted with an indifferent gas, in a rapid current of air.

(3) By means of the arc or spark discharge in air. This method which is, in principle, the same as No. (1), consists in striking an arc or spark between electrodes of any suitable metal in an air stream. Oxidisable metals readily give oxides, whilst platinum, silver, and gold yield dispersoids which are metallic. The method applied to the formation of hydrosols has been studied

by Svedberg (7), and the remarkable character of the oxides formed in air has been investigated recently by Kohlschutter and Tüscher (8).

From our point of view this method has the advantage of being easily controlled, and gives homogeneous clouds of a high degree of dispersion.

For reasons which are sufficiently obvious, it is necessary, when dealing with gaseous systems, to work with a very large volume of dispersoid. The systems studied were formed in a cubic box (fig. 1), of a capacity of 1000 litres, made of sheets of plate glass joined together by a framework of wood and lead. A lead-lined door communicated with the interior, and the whole when in order was quite air-tight. On the floor an electric fan (F) was placed and about 2 feet away from it the arc (A) or electric heater, both of which could be controlled from the outside. Fig. 1 shows a small scale diagram of these arrangements. To change the atmosphere in the box an electrically driven blower was used which pumped air into the chamber at the rate of 3 cubic feet per minute, and in order to eliminate dust and obtain an optically empty medium a filter of fine felt was interposed between the blower and the chamber.

As a guide to the type of cloud obtained by any of these methods, samples were examined ultra-microscopically with an instrument to be described later. It proved quite easy to discriminate qualitatively between coarse and fine dispersoids by noting the brightness of the particles, their Brownian movement, and their rate of subsidence under gravity. Confining our attention to aerosols of low concentration, it became obvious at an early stage that any of the above methods yield systems of considerable persistence and stability, composed of particles proved later to vary in radius from 3 to 10×10^{-6} cm., *i.e.*, comparable with those present in many colloidal solutions. Clouds of this type produce no visible fog in the chamber, but their presence can be detected by the Tyndall beam, and in some cases there is little noticeable change after 24 hours.

Determination of Number of Particles in Given Volume of Dispersoid.

Few direct methods are available for determining this important factor in gaseous systems. The only two known to us were designed for testing ordinary air, viz., the dust counter method of Aitken (9) and the striking method recently published by Owen (10).

Both these are unsuitable for aerosols, the former because the particles in our systems are so numerous that much dilution would be necessary, and the latter because the ultra-microscopic particles would probably escape detection.

We have had recourse to a direct count of the particles in the field of the ultra-microscope, and the method is simply a modification of Zsigmondy's original device used with the slit ultra-microscope for enumerating the particles in colloidal solutions. On account of the nature of aerosols, the observation cell must be of fairly large dimensions, and it is impracticable to use shallow horizontally placed cells of known depth of the kind employed for fluids with ultra-condensers of the cardioid or paraboloid types. The Jentsch type, which can be used for the observation of smokes, gives neither a sharply defined nor a narrow enough zone of illumination, and the depth of focus of the microscope objective has to be relied on to determine the volume in which the count is made. Another objection is that large particles which reflect much light remain visible as diffraction discs for considerable distances on either side of the region of sharp definition; this impairs the blackness of the background, and, while obscuring the visibility of the smaller particles, leads to an erroneous computation of the number of the larger.

The slit ultra-microscope is largely free from these defects; the delimitation of the depth of the field is accomplished by illuminating a portion of the contents of the observation cell by means of a narrow ribbon of light (the image of a brightly illuminated slit), and the thickness of this ribbon at the focal point can be determined by the simple device of rotating the slit through 90° and measuring with an eyepiece micrometer. Much depends on the source of light, its intrinsic brilliance and steadiness. After much

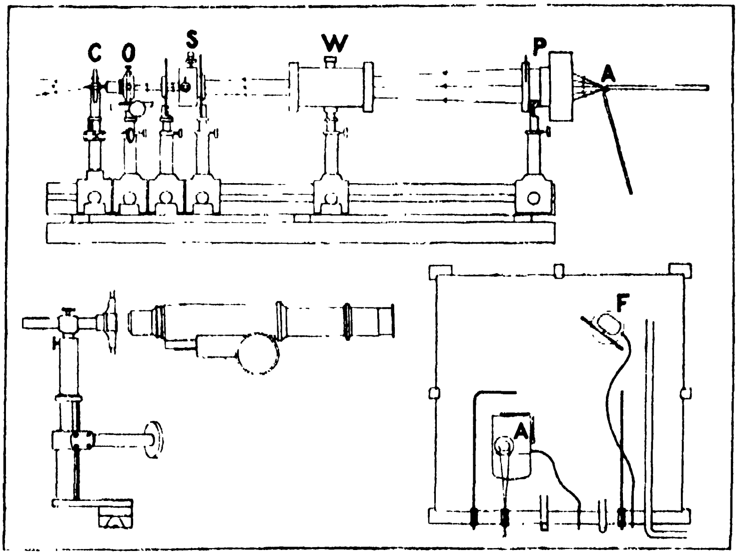


Fig. 1.

experimenting with mechanical arcs we adopted a Zeiss 5- to 10-ampère lamp, provided with a combined electromagnetic and clockwork control, which gave a very steady illumination. The light from the arc was brought to a focus on the slit, and the image of this was projected into the cell just as in the Zsigmondy instrument.

The arrangement of the various parts of the ultra-microscope is shown in the upper part of fig. 1. A is the Zeiss arc, P a parabolic condensing lens, W a water cell for removing heat rays from the illuminating beam, and S is a crossed slit adjustable both in length and breadth by fine micrometer screws. O is the projection objective, usually a 25-mm. holos, and C is the observation cell. The optical bench, focussing rackwork and other fittings, as well as the crossed slit (S), were made for us by Messrs. Males and Andrews, of Iliffe Yard, Southwark.

The observation cell (C) was easily made by cutting a small rectangular piece out of the edge of a piece of brass plate of about 7 mm. in thickness, and, after removing the superfluous metal, holes were drilled from the two ends into the aperture. After the interior had been carefully blackened, the rectangular cell was completed by cementing on to the edge and sides three windows of thin glass. The cell was mounted on the optical bench so that the light entered and escaped by the two side windows, and observations were made by means of a horizontally placed reading microscope through the third window. An enlarged diagram of the cell, mounting, and observation microscope is shown in the lower left-hand portion of fig. 1. As a rule, for counting, either a 16 or a 25 mm. Watson holos objective with a × 10 eyepiece was used. To limit the field of view for counting, a series of carefully calibrated eyepiece stops was used. Before making observations, the depth of the illuminated ribbon of light at the focal point in the cell was adjusted so as to be less than the depth of focus of the microscope, and this depth was measured just as in the Zsigmondy instrument. For setting and calibrating, dilute tobacco smoke proved convenient.

After adjusting the instrument a quick stream of dispersoid was sucked through the cell in an upward direction against gravity by means of a constant head aspirator. The flow was then checked and regulated with the help of a delicate needle valve until it was possible clearly to distinguish the particles as they slowly traversed the field seen through the eyepiece stop. Counts were then made in uninterrupted succession. As soon as the last of a group of particles had disappeared, another set was counted, until fifty fields had passed across the diaphragm. The arithmetic mean of the fifty observations was then taken and the number of particles per cubic centimetre computed. At first, any number above seven per field is difficult to count. With practice, ten, and even more can be estimated with success.

From time to time during the counting, the contents of the chamber were stirred and mixed by fanning.

To illustrate the method, the details of one of our earlier experiments are given, showing how the data were obtained and computed.

Zinc Oxide Dispersoid formed by Decomposition of Zinc Ethyl.

The depth of the light beam over the A diaphragm was 0·0031 cm. and the area was 0·00041 sq. cm. Hence the volume of the cloud viewed through the microscope was the product of these numbers, and the factor required to convert the number of particles per diaphragm into particles per cubic centimetre was $0 \cdot 786 \times 10^6$.

First Count.—Cloud dispersed at 2.00 P.M. Duration of counting 2.4–2.6 P.M.

Particles per diaphragm (50 counts).

4	5	5	7	8
5	6	8	6	7
7	8	8	7	6
6	8	4	8	8
7	6	4	6	5
8	5	8	6	7
7	4	7	5	7
8	6	8	7	8
4	6	7	6	8
8	8	6	6	8
64	62	65	64	72

Total number counted	327.
Average number per diaphragm	6·54.
Average number of particles per c.c.	$5 \cdot 14 \times 10^6$.

Ninth Count.—The diaphragm and the depth of beam were the same, and hence the same factor was required. Counting began at 3.6 P.M. and continued to 3.8 P.M.

Particles per diaphragm (50 counts).

2	1	3	3	2
3	2	1	4	2
1	1	2	4	2
3	3	2	1	2
3	1	3	1	3
2	3	3	3	3
3	3	2	2	1
1	1	3	2	2
2	2	1	1	2
1	2	2	4	2
21	19	22	25	21

Total number counted	108.
Average number per diaphragm	2·16.
Average number of particles per c.c.	$1 \cdot 7 \times 10^6$.

With this cloud twenty-four counts were made in this way and the accompanying Table shows the experimental results.

Experimental Results for Zinc Oxide.

The process of dispersal lasted 30 seconds and was finished at 2 P.M.

Mean time.	Particles per c.c.	Mean time.	Particles per c.c.
2·05 P.M.	5·14 × 10^6	3·56 P.M.	1·95 × 10^6
2·18	3·30	3·59	1·01
2·34	2·58	4·02	0·82
2·42	1·90	4·11	0·82
2·45	2·15	4·19	0·82
2·47	1·92	4·24	0·79
2·50	1·76	4·46	0·78
3·04	1·73	4·51	0·73
3·07	1·70	4·56	0·65
3·09½	1·52	5·21½	0·61
3·15½	1·45	5·31	0·61
3·18	1·46	5·40	0·54

The curve in fig. 2 was drawn from these results.

The Accuracy of the Counting.

The question as to how far the results of the counting give a true picture of the cloud as a whole is an important one.

Even the mean of fifty counts represents the particles in a very small fraction of the total cloud, and this is taken as typical of the whole dispersoid at a given period. The character of the curves, however, shows that under our experimental conditions the clouds we have examined must be approximately homogeneous in distribution, even down to comparatively small volumes, and all our experience goes to show that the method gives a fairly accurate representation of the volume distribution of these dispersoids.

Whether the values obtained for the number of particles per cubic centimetre approximate to the real number present is another question. Apart from experimental errors inherent in the slit ultra-microscope, which have often been discussed (11), there is the question of particles of amicroscopic size, beyond the limits of our instrument; but if attention is confined only to the visible particles, direct proof has been obtained in some of our earlier experiments that the numbers seen in the ultra-microscope agree closely with those deduced in another way.

It was found possible to collect the particles from a known small volume of the cloud by subsidence, illuminate them by a strong beam of reflected light in a manner described later, and then count the numbers in the deposit.

Quite a close agreement was obtained by the two methods, as the following Table shows, but it is to be noted that the particles were relatively of a large size, of the order of 10×10^{-6} cm. radius. Further pointolite illumination was employed in both cases.

Time in minutes.	Particles per c.c.	
	From deposit.	From ultra-microscope.
25	1028×10^3	960×10^3
45	701	680
75	335	350
105	241	280
135	243	220
207	141	141
255	98	91
495	57	56

It is not easy to deal with much finer particles in this way. Using arc illumination and a very highly dispersed cloud of cadmium oxide, we have attempted to check our more recent figures; but on account of the difficulty of seeing the particles once they have settled, and also because a satisfactory surface is difficult to obtain, the results have been disappointing.

But while the values obtained for the number of visible particles per cubic centimetre may be approximately correct only, there is little doubt that counting with the ultra-microscope enables a relative comparison of different aerosols to be made with some accuracy, once experience has been gained.

It is obvious from the figures just cited that the number of particles in a cloud diminishes rapidly with time. It is quite impossible that this fall in number can be due either to subsidence under gravity on the floor or deposition by Brownian movement on the walls of the chamber. Calculation and experiment alike show that in a cubic-metre box the number of particles removed in this way is a small fraction only of those found to disappear. This rapid change in a freshly formed smoke seems a characteristic feature of gaseously dispersed matter.

Although we have examined a large number of clouds composed of substances varying greatly in chemical and physical properties, they all exhibit at the start this remarkable behaviour.

These changes can best be followed from the curves obtained by plotting the number of particles per cubic centimetre against the age of the cloud, reckoned from the time when dispersal is completed. For convenience they may be called "particulate number curves." Several of such curves are given in the diagram (fig. 2), and Table I shows the data from which they have been drawn.

Curves for different substances vary considerably in the number of particles present at corresponding times. Curves for the same substance only approximate closely when the method and conditions of formation are closely similar.

All curves dispersed in pure dust-free air exhibit certain characteristics in common, which in a sense define the life-history of these systems. The changes fall into three periods, viz.:—

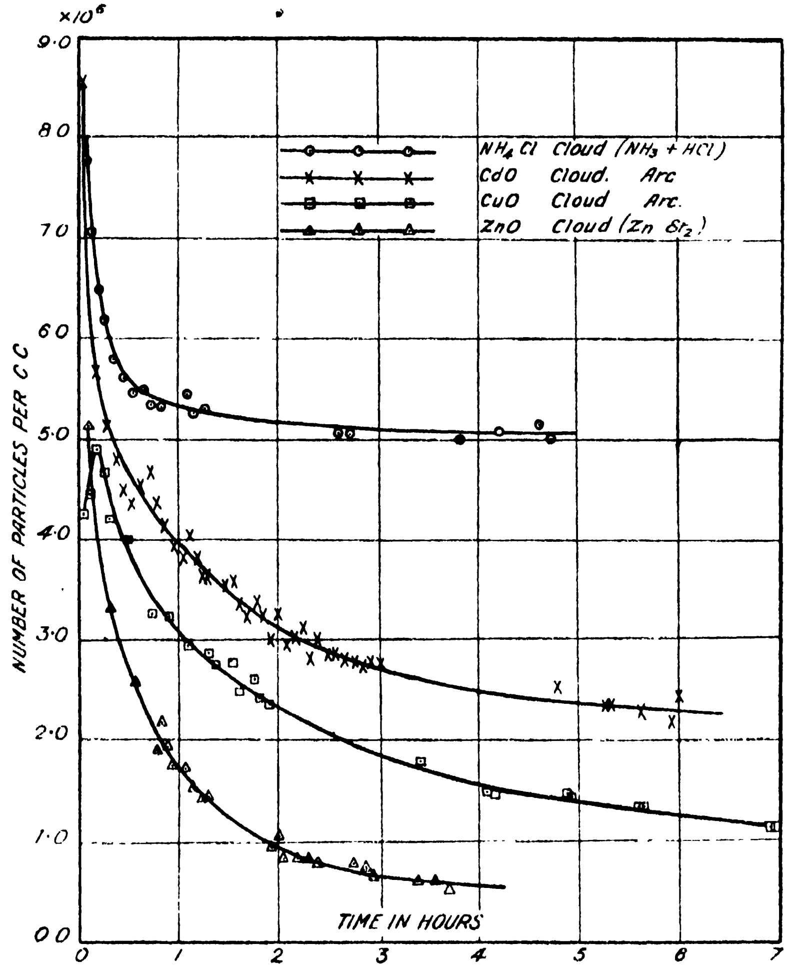


Fig. 2.

(*a*) An unstable initial period, in which the decrease in the number of particles with time is very rapid.

(*b*) A stable period, reached only after five hours or longer, in which the change in particulate number with time is very slow, and continuing until the cloud has settled out on the walls or floor of the chamber or dissipated from other causes. This stage often lasts for twenty-four hours or longer even in a chamber of 1 cubic metre capacity.

(*c*) An intermediate stage, where the rate of disappearance is the resultant of the factors operative in (*a*) and (*b*).

Table I.

Time.	Particles per c.c.	Time.	Particles per c.c.	Time.	Particles per c.c.
Cadmium Oxide Cloud.—Number Curve.					
3 m.	8·54 × 10^6	1 h. 14 m.	3·62 × 10^6	2 h. 24 m.	3·00 × 10^6
10	5·65	1 16	3·63	2 29	2·84
16	5·10	1 28	3·54	2 34	2·86
22	4·80	1 33	3·58	2 40	2·80
26	4·48	1 37	3·35	2 46	2·78
32	4·35	1 41	3·21	2 50	2·71
36	4·54	1 47	3·39	2 53	2·77
41	4·66	1 51	3·23	3 1	2·75
46	4·37	1 57	3·00	4 48	2·54
52	4·11	2 1	3·26	5 17	2·34
58	3·93	2 6	2·95	5 19	2·35
1 h. 4	3·82	2 10	3·02	5 37	2·29
1 8	4·03	2 15	3·11	5 56	2·17
1 12	3·81	2 19	2·80	6 1	2·44
Copper Oxide Cloud.—Number Curve.					
2 m.	4·25 × 10^6	1 h. 19 m.	2·88 × 10^6	4 h. 4 m.	1·51 × 10^6
6	4·45	1 21	2·66	4 8	1·47
10	4·92	1 33	2·70	4 53	1·48
15	4·67	1 35	2·49	4 55	1·45
19	4·20	1 45	2·60	5 35	1·35
29	3·99	1 47	2·48	5 37	1·35
44	3·27	1 53	2·34	6 56	1·12
55	3·22	1 54	2·26	6 58	1·16
1 h. 6	2·94	3 24	1·80		
Ammonium Chloride Cloud.—Number Curve.					
4 m.	7·76 × 10^6	31 m.	5·47 × 10^6	1 h. 14 m.	5·30 × 10^6
7	7·06	38	5·50	2 37	5·07
11	6·48	42	5·35	2 44	5·06
15	6·19	49	5·34	3 48	5·00
21	5·79	1 h. 5	5·43	4 12	5·09
26	5·61	1 9	5·28	4 24	5·15
				4 28	5·01

Now these curves show a close resemblance in form to those obtained by Tolman and his colleagues with dense smokes of various materials (oil, rosin, acetanilide, and benzoic acid) by tracing the changes by means of the variation in reflected and scattered light with the help of the Tyndallmeter (12). This method gives a measure only of the combined effect of a change both in number and size, but Tolman was able to correlate both factors independently with the observed intensities, and to follow by indirect methods the separate effect on the readings of his instrument of these two factors.

He concluded, after much careful investigation, that the disappearance of smoke in a closed space was brought about mainly by the aggregation of the particles themselves, and was only influenced to a minor degree by adhesion to the walls or by settling out.

The initial concentrations used were of the order of 800 mgrm. per cubic metre and the smokes had the appearance of dense fogs, so that it is not surprising that they proved very unstable. It is very probable, too, that the particles which radiated most of the light were comparable with the wave-length and much larger than ours, but it is interesting that the changes in these relatively coarse and highly concentrated dispersoids should be so similar to those in the dilute systems we have been studying.

Growth of the Particles.

When viewed through the ultra-microscope the particles present during the period of quick change are seen to increase rapidly in size. As a rule, immediately after formation, the Brownian movement in these clouds is very marked, but the particles scatter relatively little light and their diffraction discs are difficult to see, but soon the brisk agitation of molecular bombardment slackens and, as the numbers diminish, brighter discs make their appearance and the dispersoid is seen to consist of particles of various sizes.

After some time the bright particles appear to predominate, and sometimes reach such a size that they reflect light of the characteristic colour of the substance *en masse*, and their brilliant luminosity in the restricted field of the counting diaphragm makes it frequently a matter of extreme difficulty to perceive the very small particles which often accompany them.

It is clear that the unstable initial period in these systems is associated with the growth of the particles. Now as long as we confined our attention to systems of organic substances possessing at ordinary temperatures even a minute volatility, it was difficult to reach any definite conclusion as to the process of growth. Two hypotheses explained equally well the increase in size and the concomitant decrease in numbers. On the one hand, the process might be attributed to the gradual aggregation of the smaller particles to form larger complexes, or, on the other, to the isothermal distillation from the smaller to the larger, and from the uncharged to the charged. In support of the latter idea there are the observations both of Millikan (13), and of Ehrenhaft (14), that charged droplets of mercury suspended in an electric field perceptibly diminish in mass in a short time, and we ourselves have noticed the complete disappearance by evaporation of the particles in a dilute cloud of diphenylamine, though the vapour pressure of this substance at ordinary temperatures is of the order of

0·001 mm. of mercury. Again, actual coalescence of particles in the field of the ultra-microscope had not been observed, and the larger and brighter particles could never, at this stage of the investigation, be resolved into complexes, but retained the appearance of brilliantly luminous spheres even when sharply focussed. The evidence, such as it was, favoured the idea of evaporation and condensation.

The later study of disperse systems of metallic oxides possessing an extremely minute and much smaller volatility changed our views entirely. It was noticed first that the particulate number curves for these systems were very similar in form to those of the more volatile organic substances pointing to the probability of the same process predominating in each.

Then, by the use of objectives of higher optical aperture combined with a × 20 eyepiece, indications of the complex character of the brighter particles were obtained. Finally, all doubts were set at rest when the aerosols of CdO and ZnO were examined. The ultra-microscopic particles of these two substances do not form compact complexes, but unite to minute strings and chain-like aggregates in which the diffraction images of the individual particles are clearly visible. When a cloud of either is produced so as to form a dispersoid of a concentration of about 20 mgrm. per cubic metre, the particles at first are in lively Brownian movement and show no sign of complexity, but in a short time the motion becomes slower, and soon, on careful focussing, little flexible chains of irregular conformation make their appearance.

These aggregates continue to grow, forming often branching structures in which the tendency to unite in strings is well marked. The behaviour of these complexes under molecular bombardment is remarkable; they twist and whirl about in a striking manner, showing clearly the flexibility of the chains of which the particles constitute the links. Sometimes the stringy complexes are seen to be in a state of internal vibration, the individual particles showing a kind of independent restricted Brownian motion which has not been described before, but recalling McBain's (15) description of the motion of particles up and down a curd fibre. In an electrostatic field the complexes straighten out and arrange themselves parallel to the lines of force, and on reversal of the field rotate through 180°. This behaviour could be explained if the chain consisted of a number of oppositely charged particles separated from each other by an insulator, which in this case may well be a film of air. The strings would then behave like a chain of small magnets in a magnetic field.

When the dispersoid is collected by exposing glass slides to the cloud and the deposits illuminated by a concentrated beam of reflected light from an arc, the aggregates can be seen more clearly and disclose the same structure. Each chain appears like beads on a thread, separated from each other by a small space.

When it is remembered that the beads are really diffraction images formed by much smaller particles, it is obvious that the latter must be some distance from each other and must be held together by invisible threads, consisting probably of strings of amicrons or invisible aggregates.

Another possibility is that the particles are joined together by a layer of adsorbed air, and that when two particles collide the cushions of air surrounding each particle coalesce, binding them together. That the particles in these clouds are surrounded by an adsorbed layer is inherently probable, and there is much evidence in the literature to support this, but the air film cannot be responsible for the chain-like structure, and the formation of the strings and threads must be due to the polar character of the particles of zinc and cadmium oxides.

Neither is the structure primarily due to the charges carried by the particles, for we have obtained exactly the same type of complex with zinc oxide obtained by the decomposition of zinc ethyl with air, a cloud which contains very few charged particles.

When much denser clouds are produced initially by raising the concentration of the dispersoid, so as to form a fog in the chamber, the particles agglomerate rapidly to large and irregular masses. When these are examined by reflected light the same structure is visible, and it is clear that they are built up of chains and branching threads arranged in an irregular fashion to form loose and porous agglomerates of a snowflake-like character. These deposits adhere strongly to the glass slides, but many hairs are usually visible which are attached only at their lower ends, and in a current of air these can be blown about, behaving like flexible but elastic fibres. By transmitted light many of the finer fibres or threads are invisible, but that they are connected with each other can easily be demonstrated by allowing a drop of immersion oil to flow gently across the deposit in the field of the microscope. The visible particles, as they are lifted up by surface tension, are seen to be attached by invisible threads to constellations of others and drag these with them from a considerable distance in front of the advancing oil. When wetted by the oil the complexes break up immediately. The radii of the small particles composing the threads are of the order of $50\,\mu\mu$, or 5×10^{-6} cm., and the larger complexes are often $30\,\mu$ in length, or longer.

The aggregate of both zinc and cadmium oxides are similar, but the tendency to string formation is more pronounced in the case of cadmium oxide, and on some occasions very long threads or chains have been noticed.

Kohlschutter and Tüscher (*loc. cit.*), in a very interesting paper on the precipitation of oxides dispersed from various metals by means of the arc discharge in air, describe the properties in bulk of the different substances collected in their electrical precipitator, and they call attention to a curious effect often observed with CdO and never noticed with any other oxide, viz., the sudden fall of potential in their precipitator, an effect which they attribute to the formation of string-like complexes of high conductivity, stretching from the earthed cylinder to the central negatively charged wire. They observed also a similar effect when the dispersed oxides of zinc and tin were electrified by means of a point discharge and then passed through a parallel plate condenser; long threads were observed to form on the earthed plate parallel to the lines of force and to grow by attracting fresh particles. Other disperse oxides showed quite a different structure. This is quite in conformity with our experience, and the character of the oxides collected in bulk is probably quite similar to that of the aggregates just described.

We have also examined the deposits obtained from other oxide clouds. The tendency to form string-like complexes seems weakest with the oxides of Pb, Cu, Mn and Cr. It is slightly greater with iron (17), whilst Mg, Al and Sb resemble zinc, and form easily loose flock-like aggregates. The structure of the complex is dependent clearly on the nature of the substance of which the individual particles are composed.

Prof. McBain (18), in a letter to 'Nature,' drew our attention to the properties of the cadmium alcosol, discovered by Svedberg (19), (and described in the report of a general discussion on colloids held by the Faraday Society in October, 1920). This remarkable sol, obtained by electrical pulverisation, consists of metallic ultramicrons of radius 5 $\mu\mu$, which on slight oxidation, and on standing, sets to a firm gel, so that the bottle can be inverted, but, on stirring, the viscosity falls again at once nearly to that of pure alcohol. This is no doubt due to the formation of a structure such as we have described, throughout the liquid, the chains and threads linking up together to form an irregular network. Indeed, it seems likely that our thicker deposits of ZnO and CdO possess a conformation closely analogous to many gels, and the term "aerogel" suggested by Prof. Donnan is not inappropriate.

The structures we have described are clearly visible to the eye through the microscope when a suitable type of illumination is used, but they are not so easy to photograph. The prints reproduced in this paper are from negatives taken with a 4-mm. objective, and the deposit was illuminated by a narrow pencil of light passing below the objective, and brought to a focus on the slide at an angle of incidence of about 30°. The source of illumination was a mercury vapour lamp, and the light was rendered approximately monochromatic by passage through a Wratten blue filter. The larger complexes yielded good photographs, but the smaller aggregates were very difficult to reproduce. The main features, however, are to be seen in the illustrations (Plate 5), which are zinc oxide complexes, each slide exposed to the cloud for ten minutes.

Having thus shown that the growth in size and decrease in numbers of the particles in these oxide-clouds is due to aggregation, it remains to explain why the process of collision is not apparent when these dispersoids are viewed with the ultramicroscope. Anyone who has observed the movements of these particles cannot fail to be struck by the way they approach, rotate about each other, and separate again without sticking.

The appearance of the field is, however, deceptive, for the diffraction discs are many times greater than the particles themselves, and hence, after collision, the doublets formed will have much the same appearance as before, exhibiting only an increased luminosity.

Further, unless collision takes place in the focal plane of the microscope, the detection of coalescence is difficult, and as the particles are moving irregularly in three dimensions, their disappearance in most cases is due simply to their going out of focus or out of the illuminated zone. Finally, it is easy to calculate from the slope of the particulate number curves that the number of collisions to be expected, even when aggregation is taking place most rapidly, is very small in the volume of dispersoid under observation. For instance, using the largest diaphragm, which takes in about a third of whole field of the microscope, and which has an area of 0·0059 cm.^{2}, the number of expected collisions would only be of the order of one per minute with a light beam of average depth, illuminating a disperse system containing per c.c. about 7 million particles. After much searching, we have at last been able to distinguish these occasional encounters between particles.

REFERENCES.

(1) de Broglie, 'C. R.,' vol. 148, pp. 1315–1318 (1909).

(2) Ehrenhaft, 'Physik. Zeitsch.,' vol. 10, p. 308 (1909).

(3) Millikan, 'The Electron' (Univ. of Chicago Press), (1917).

(4) de Broglie and Brizard, 'C. R.,' vol. 148, p. 1457 ; vol. 149, p. 923 (1909).

(5) 'Brit. Assoc. 4th Report on Colloid Chemistry,' p. 122.

(6) Tolman and Vliet, 'J. Amer. C. S.,' vol. 41, pp. 297–300 (1919) ; Tolman, Reyerson, Vliet, Gorke, Brookes, 'J. Amer. C. S.,' vol. 41, p. 300 ; also p. 575 (1919).

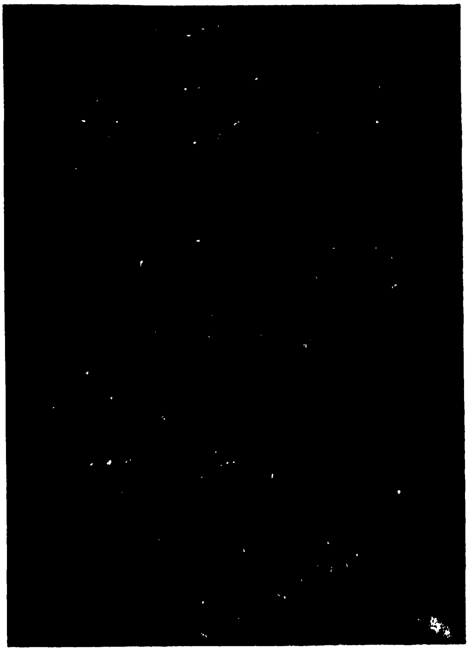
(7) Svedberg, 'The Formation of Colloids' (Churchill).

(8) Kohlschutter and Tüscher, 'Zeitsch. f. Electrochem.,' vol. 27, p. 1 (1921).

(9) J. Aitken, 'Roy. Soc. Proc. Edin.,' vol. 16, p. 167 (1890).

(10) J. Aitken, 'Trans. Roy. Soc. Edin.,' vol. 30, p. 337 (1883) ; Owen, 'Roy. Soc. Proc.,' A, vol. 101, p. 18 (1922).

I. 9 minutes exposure.



II. 29 minutes exposure.

I.—49 minutes exposure.

IV.—244 minutes exposure.

(11) Zsigmondy, 'Zeitsch. f. Physik. Chem.,' vol. 92, pp. 619-622 (1918).
(12) Tolman and Vliet, *loc. cit.* (6).
(13) Millikan, 'Phys. Rev.,' vol. 32, p. 389 (1911), and 'The Electron,' p. 168.
(14) Schlidlof and Karpowicz, 'C. R.,' vol. 158, p. 1912 (1914); Ehrenhaft, 'Phys. Zeitschrift,' vol. 16, p. 227 (1915).
(15) J. W. McBain, 'Roy. Soc. Proc.,' A, vol. 98, p. 395 (1921); 'J. Chem. Soc.,' vol. 117, p. 1506 (1920).
(16) Kohlschutter and Tüscher, *loc. cit.*
(17) Elihu Thomson, 'Nature,' p. 619 (July 14, 1921).
(18) McBain, 'Nature,' p. 683 (July 28, 1921).
(19) Svedberg, 'Discussion on Colloids,' Faraday Society, October, 1920.

Smokes: Part II.—*A Method of Determining the Size of the Particles in Smokes.*

By R. Whytlaw-Gray and J. B. Speakman.

(Communicated by Prof. A. Smithells, F.R.S. Received October 25, 1922.)

In Part I (1) an ultra-microscopic method of counting the particles in dilute and highly disperse solid-gas systems was described, and whilst the number was found to decrease as the cloud aged, a concurrent growth in size of the individual units was observed to take place. These changes in the case of non-volatile oxide smokes were proved to be caused by the aggregating of the smaller particles to form larger complexes.

Now, in order to follow this process in a quantitative manner, it is necessary to determine the average mass of a particle at different periods in the age of the smoke. That is to say, some method must be used which can be carried out quickly, so that the particles have not time to increase appreciably in size. Now, none of the usual methods which deal with the movement of individual particles are suitable, because, apart from the validity of the Stokes-Cunningham equation (2), on which the majority of them are based, they are in practice too slow, and require the mean of a large number of observations if a representative value for the average size is to be obtained. This objection applies even more strongly to methods based on Einstein's (3) equation of the mean displacement of a particle by molecular bombardment along a given axis in a definite time.

Again, Millikan's (4) balancing drop method, which is independent of Stokes' law, could hardly be applied to particles so small as ours, which exhibit so rapid a Brownian movement. If we could assume the Stokes-

Cunningham law to be applicable for these very small particles, and further be sure that in these clouds they all carried equal electric charges, then the ingenious method of Wells and Gerke (5) might be applied, for an average value can be obtained quickly; but these systems contain always a large number of uncharged particles, and the equality of charge on the others is by no means certain.

Experimental.

These considerations induced us to work out a method adapted to this particular problem, and which has the advantage of directness.

Since we had a large volume of dilute smoke to deal with, it seemed possible to determine by filtration the total weight in a given volume and deduce the mean mass per particle by counting simultaneously the number in the same volume. In practice it was found possible to carry this out by sucking a measured volume of cloud, usually 1 litre, through a small asbestos filter, and determining its increase in weight on a micro-balance.

A good deal of experimental work was necessary before a really satisfactory filter was devised. At first, carefully levigated white Gooch asbestos was packed in the moist state into thin-walled glass tubes about 2 cm. in length and about 3 mm. internal diameter. These were constricted at a point about 7 mm. from one end, thus providing a seating for the asbestos pad (fig. 1).

It was found, however, that loss of asbestos was liable to occur, and concordant results were difficult to obtain. Moreover, when tested with the ultramicroscope, some of the visible particles were seen to penetrate the filters, especially in the early stages, and with certain types of dispersoid.

The difficulty of the loss of asbestos was surmounted by treating the pads *in situ*, after they had become thoroughly dried, with a weak solution of collodion in acetone, and then drying them again in the oven. This treatment had the desired effect of binding together the loose fibres, and so preventing their removal, when a quick stream of cloud was passed through the filter.

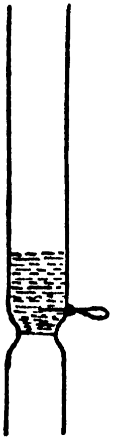


FIG. 1.—Microfilter.

The passage of the particles through the filter in the early stages of the cloud was more difficult to prevent; various materials were tried, but none was found to be superior to the asbestos. After a good deal of experimenting, the device was adopted of sealing a small platinum point through the side of the tube, so that it projected a small distance into the asbestos, and connecting it when filtering the cloud with a small Wimshurst machine.

In this way a point discharge was produced in the filtering layer, and the little tube functioned as a combined filter and electrical precipitator. This arrangement proved perfectly successful in stopping all the visible particles when used under the right conditions.

Even with these improvements, the results in a series were liable to show sometimes unaccountable variations. The efficiency of the asbestos depends on the velocity of flow being sufficient to ensure the requisite degree of turbulence when the dispersoid is being sucked through it. The size of the pores in the material is of course very much greater than the diameter of the largest particle, and it has been noticed that increasing the length of the filtering material often makes the process less efficient, on account of the diminished rate of flow caused by the increased resistance. To get comparable and concordant results the filters should be as nearly as possible of the same resistance, so that with a constant difference of pressure between the two sides, the times taken for the filtration of equal volumes of dispersoid are nearly the same. In carrying out a determination, the filter was connected on the one hand with the chamber containing the dispersoid, and on the other with an aspirator and pressure gauge; the tap was then opened and adjusted until the pressure in the aspirator was about 24 inches of water below atmospheric. The outflowing water was collected and measured, which gave after correction the volume of dispersoid filtered. Usually, a volume of 1 litre was filtered for each determination, and the process lasted for about 5 minutes.

In the majority of the clouds we have investigated the weight concentration was of the order of 20 mgrm. per cubic metre, so that the increase in weight of the filter was about 0·020 mgrm. The micro-balance employed was of the type used by Steele and Grant (6), and by Ramsay and Gray (7), and whilst carrying a load of about 200 mgrm. was sensitive to 0·0002 mgrm., an accuracy amply sufficient for our purpose. In determining this small difference in weight, it was found quicker and more accurate to keep the air pressure in the case at a constant and low value, and to rely on the change in zero of the balance to measure the weight of the particles caught.

This eliminated the use of the small bulb of known volume, except for the purpose of calibrating in terms of weight the shift of the beam. The zero displacement was observed and measured by means of a travelling reading microscope moved by a micrometer screw, and on calibration with the buoyancy bulb it proved to be directly proportional to the weight applied, as can be seen from the accompanying figures.

Calibration of Microbalance.

Temperature.	Pressure in case.	Micrometer reading.
° C.	mm. Hg.	mm.
16·4	48·8	11·072
16·2	176·7	9·414
16·5	244·6	8·528
16·6	279·2	8·042
16·6	328·8	7·356
17·0	360·9	6·835
16·9	93·9	10·480
17·0	30·9	11·295
17·2	60·7	10·921

The volume of the bulb was 158·04 c.mm. Hence from interpolation each millimetre displacement of the zero represented a change in weight of 0·01914 mgrm., and since the drumhead of the micrometer was divided into 100 large divisions, each of these was equivalent to a change in weight of 0·0001914 mgrm.

In making a series of measurements on a cloud the method of procedure was as follows:—A number of small tubes to contain the filtering material were first made and their weight adjusted to approximate equality. This was easily done by grinding down the ends of the heavier ones on a fine carborundum stone until no difference in weight could be detected on an ordinary balance. The tubes were then filled with asbestos, and after treatment with collodion were tested with the aspirator and the gauge to see if their resistances were nearly the same. Any abnormal filters were rejected, and the remainder were made very nearly equal in weight by the addition of minute fragments of asbestos. Finally, the filters were carefully cleaned, weighed, and preserved in a desiccator until required. The results obtained in this way yielded a series of values for the weight of the visible particles in a litre of dispersoid at various intervals during its life-history. Curves were obtained by plotting the weight per litre against the age of the cloud. In the majority of cases the graphs for different substances, and for clouds of the same substance studied on different occasions, proved to be very similar in form, as fig. 2 shows. The micro-balance results obtained from four different clouds are shown below.

It is now possible, by combining the results of the counting with the determinations of weight, to find the average weight of a particle in the cloud at different periods. Thus, knowing the total weight per litre of the dispersoid, and dividing this by the number in the same volume at the same period read from the number curves, the average mass per particle is obtained in a direct

Microbalance Curves.

Time.		Weight (mgrm. per litre).	Time.		Weight (mgrm. per litre).
Cadmium Oxide Cloud 1.			**Cadmium Oxide Cloud 2.**		
h.	m.		h.	m.	
0	7½	0·0208	0	5	0·0094
0	36	0·0212	0	24	0·0151
1	46	0·0179	0	54	0·0138
3	42	0·0113	1	25	0·0125
4	17	0·0078	2	45	0·0109
4	29	0·0074	4	20	0·0086
5	39	0·0048	4	33	0·0097
			7	20	0·0080
Zinc Oxide.			**Copper Oxide.**		
h.	m.		h.	m.	
0	24	0·0220	0	24	0·0043
1	1	0·0300	0	38	0·0051
1	28	0·0280	1	6	0·0048
1	55	0·0250	1	28	0·0043
2	53	0·0197	3	16	0·0007
3	45	0·0171	3	32	0·0004
4	9	0·0158	3	47	0·0013

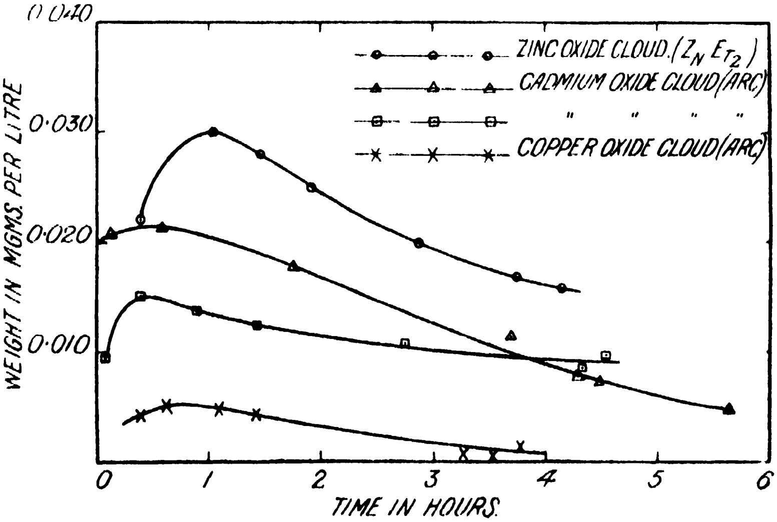


FIG. 2.

manner. To obtain reliable values it is, of course, necessary to determine weight and number on the same cloud.

We have made a number of parallel determinations of this kind on various types of dispersoids. The results cited below were obtained with zinc oxide and are typical.

Zinc Oxide Cloud.—Number Curve.

Time.	Particles per c.c.	Time.	Particles per c.c.	Time.	Particles per c.c.
h. m.		h. m.		h. m.	
0 4	$4·54 \times 10^6$	1 51	$1·68 \times 10^6$	2 47	$0·88 \times 10^6$
0 26	2·99	1 53	1·42	2 49	0·77
0 29	2·89	1 55	1·31	2 51	0·69
0 51	2·09	1 58	1·59	4 7	0·56
0 53	2·25	2 28	0·91	4 10	0·48
1 21	1·68	2 33	0·81	4 16	0·56
1 24	1·65	2 34	0·92		

Zinc Oxide Cloud.—Microbalance Curve.

Time.	Weight (mgrm. per litre).	Time.	Weight (mgrm. per litre).
h. m.		h. m.	
0 10	0·0172	2 46	0·0124
0 35	0·0187	4 23	0·0081
1 3	0·0171	4 36	0·0079
1 33	0·0141	4 50	0·0076
2 11	0·0147		

The graphs of these experimental values are shown in figs. 3 and 4 (*a*). In calculating the average mass per particle, the smoothed values for weight and number were taken from the curves at corresponding times, and the results are tabulated below:—

Table showing the Average Weight of a Particle at different Periods.

Time.	Average weight per particle, grm.	Time.	Average weight per particle, grm.
h. m.		h. m.	
0 4	$3·62 \times 10^{-15}$	3 0	$15·1 \times 10^{-15}$
0 30	6·60	3 30	16·4
1 0	8·37	4 0	16·9
1 30	9·82	4 30	16·9
2 0	11·7	5 0	16·8
2 30	13·6		

When these figures are graphed the curve in fig. 4 (*b*) is obtained, which demonstrates the growth of the particle up to about 4 hours from the time of dispersal of the cloud. The average mass then becomes constant, at a value which is about 4·7 times its initial mass.

If the assumption is made that the particles are compact spheres of density 5·6 (the density of zinc oxide in bulk), the average radius can be evaluated.

It has been shown, however, that even the smallest particles consist in all probability of aggregates, so that the radii calculated give only a rough measure of the true size.

If we take the particles of smallest average mass, *i.e.*, $3{\cdot}6 \times 10^{-15}$ grm., the

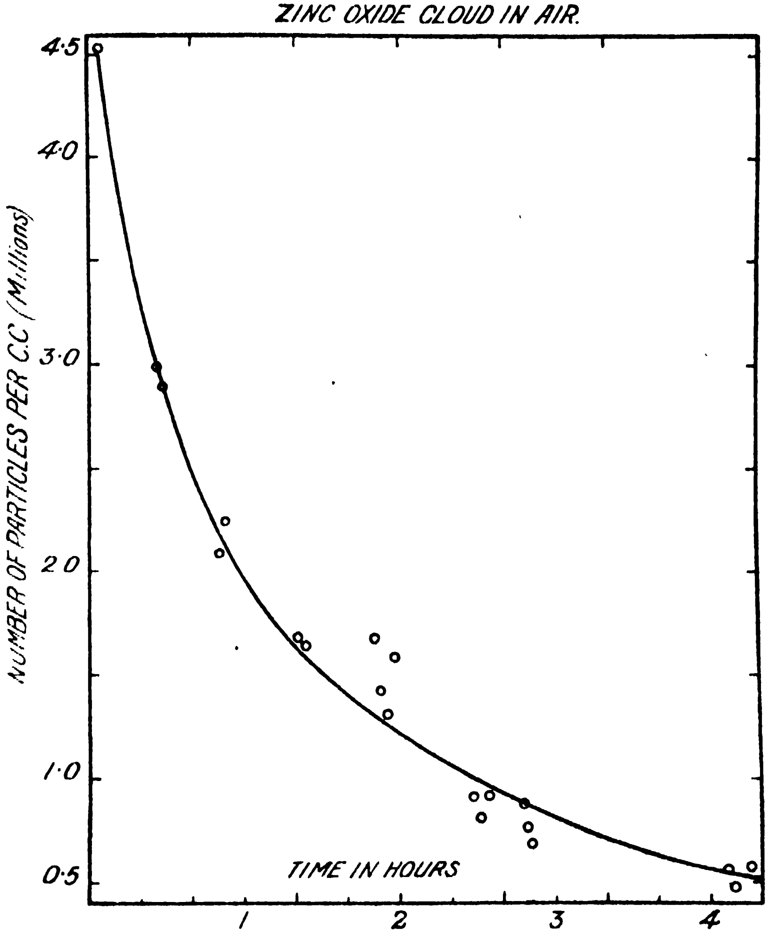


FIG. 3.

corresponding radius is $5{\cdot}4 \times 10^{-6}$ cm., or 54 $\mu\mu$. About 50 $\mu\mu$ seems to be a usual value in the initial stages of many gaseous dispersoids, and we have obtained it with many different substances. It must be remembered, however, that this is an average value and that there are many smaller particles present.

We have made a number of experiments to find what the radius of the smallest particle is which is visible in our ultra-microscope. It is not difficult when working with the arc to produce oxide clouds of varying degree of dispersion by changing the duration of the discharge and the velocity of mixing. It has also been found possible to remove by electrical methods the larger particles in a cloud and then determine the mass of those that remain. We have succeeded in this way in obtaining cadmium oxide dispersoids consisting of particles of radius 21 $\mu\mu$, or $2{\cdot}1 \times 10^{-6}$ cm., and on one occasion of 15 $\mu\mu$, or $1{\cdot}5 \times 10^{-6}$ cm. Since these are average values there is little doubt that particles as small as 10 $\mu\mu$, or $1{\cdot}0 \times 10^{-6}$ cm., are clearly visible—the same order of magnitude as the particles in a fine gold hydrosol. As might be expected from the type of ultra-microscope necessary for this class of work, we have not been able to reach the same limit of visibility as Zsigmondy has attained with metallic hydrosols; for using his latest type of instrument, he has found it possible to count particles of as small a radius as 3 $\mu\mu$, or $0{\cdot}3 \times 10^{-6}$ cm.

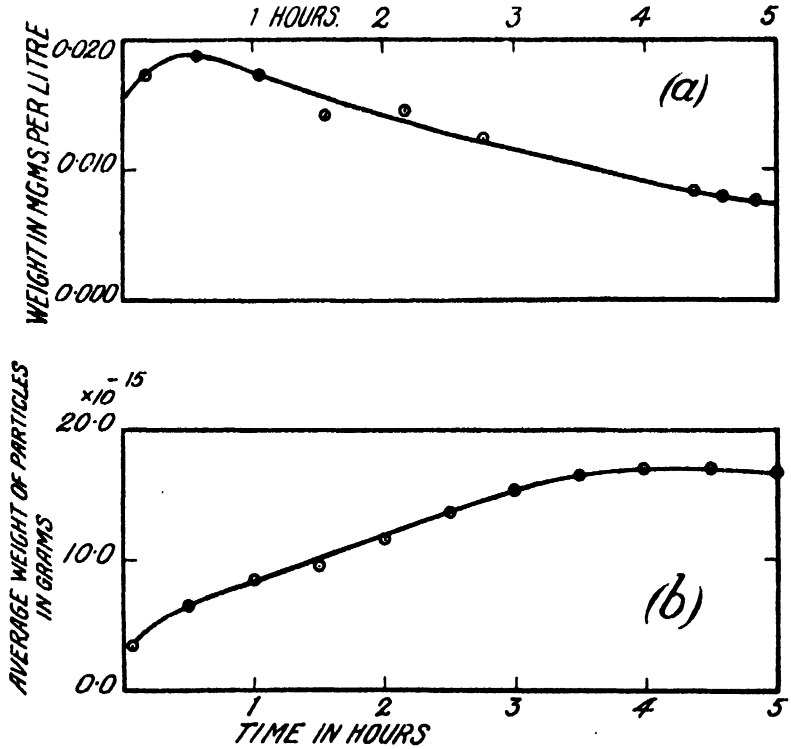


Fig. 4.

Returning to the micro-balance curves in figs. 2 and 4 (*a*), they all exhibit one remarkable feature which was quite unexpected. The weight concentration increases in nearly all cases during the early stages of a cloud, when the number of particles is diminishing rapidly and all the curves show a well-defined maximum. It would not have been surprising if the weight had remained constant during the period of quick change, for that would have indicated that the fall in particulate number was due to aggregation or coalescence. That the weight should rise initially can only mean that the dispersoid contains at the start either (*a*) molecularly disperse substance, which is condensing around the particles present, or (*b*) large numbers of amicrons, invisible in the ultra-microscope, which pass through the filter. These amicrons would in time either coalesce to form visible and filtrable aggregates or they would attach themselves to the larger particles and so cause an increase in weight of the dispersoid caught. Now a glance at the diagram shows that this increase in weight continues from 30 to 60 minutes after the cloud has been dispersed, and it is difficult to believe that such a high degree of supersaturation as this would involve can exist for so long in presence of the cloud itself, which contains upwards of 1×10^6 particles per cubic centimetre to act as centres of condensation. Again, molecules so far below their condensing temperature would almost certainly stick to any surface they hit (8) and would in all probability condense in passing through the micro-filter. Amicrons, on the other hand, might be expected to penetrate the filters, especially when the difficulty of catching the finer visible particles is recalled.

Apart from the micro-balance curves, direct evidence of the presence of amicroscopic or molecular matter has been obtained by filtering the dispersoid in its early stages and examining the medium from time to time. To do this either the micro-filters themselves were used, or a larger filter of the same type, and immediately after dispersal a few litres of the cloud were filtered into a clean dry flask containing dust-free air, or in some of our recent experiments, into an evacuated carboy, care being taken during the process to maintain the fall of pressure across the filter at a normal value. When filtering into the empty carboy, this was accomplished by the adjustment of a screw-down needle-valve. In all cases a positive result was obtained, and the filtered air, either optically empty at first or containing relatively few ultra-microns, was seen to become full of visible particles after a period which varied from 30 minutes to 2 hours. By counting the particles it is possible to measure the rate at which they grow. A fuller study of this interesting phenomenon is being undertaken and will form the subject of a third paper. It is probable that what is taking place among the amicroscopic particles is similar to, if not identical with, the changes among the visible particles during the unstable period.

Now if we assume the invisible portion of the cloud to be particulate and not molecular, which for reasons just stated is a very probable assumption then it is possible to get a rough idea of the number of the amicrons

Taking the observed rise in the micro-balance curve for zinc oxide (fig. 2), 0·008 mgrm. per litre, and supposing that the average radius of the amicron is 6×10^{-7} cm., a value slightly below the limiting size visible in our instrument, *i.e.*, 10×10^{-7} cm., the number per cubic centimetre must be $1·6 \times 10^{9}$. Even if we suppose the radius to be 10×10^{-7} cm. the number per cubic centimetre comes out to $3·4 \times 10^{8}$.

It is hence clear that if amicrons are present at all they must far outnumber the ultra-microns, and the behaviour of these systems, in general, may be influenced largely by the presence of this invisible material.

Having shown that this method of finding the average mass is a reliable one, it seemed desirable to attempt a more quantitative study of the process of aggregation to discover whether the formation of complexes proceeded approximately at the same rate as the disappearance of the particles. The idea was that if the average number of particles per complex could be estimated, then the product of the degree of complexity and the particulate number should be constant throughout the life of the dispersoid. This procedure cannot be expected to give a result unless allowance is made for the number of particles caught on the walls and floor of the chamber during the period of the experiment. The fall in weight concentration determined by the micro-balance furnishes, however, a measure of this and enables the calculation to be made. The actual method was as follows:—

To find the average degree of complexity eight carefully polished glass slides were exposed successively, for 30 minutes each, to a zinc oxide cloud of an initial concentration of about 20 mgrm. per cubic metre.

Zinc Oxide Cloud.

Number of slide.	Mean time from dispersal.	Mean number of particles per complex.
	h. m.	
1	0 25	1·20
2	1 0	1·37
3	1 35	1·58
4	2 2	2·00
5	2 40	2·69
6	3 12	2·70
7	3 45	3·33
8	4 17	3·61

Each slide was allowed to remain in the cloud in a horizontal position for 30 minutes and then removed. The deposits were then examined, using for illumination a strong beam of reflected light, and the relative number of one's, two's, three's, etc., present on the average in a field were tabulated. From these observations the mean number of particles per complex was computed, a number which gives a measure of the average degree of aggregation reached by the cloud at any interval after its formation.

These numbers are roughly approximate only. They do not give a true picture of the frequency of occurrence in the cloud of aggregates of various degrees of complexity, but only the sedimentation value, and from the structure of the particles and their aggregates it is difficult to get any idea of the relative rate at which they deposit. It is to be noted, too, that the " singles " are not primary particles and are almost certainly aggregates. Again, the presence of convection currents in the chamber probably alters the proportion of the various complexes. But if these errors are disregarded it is possible to see whether the fall in numbers is accounted for roughly by the formation of aggregates.

Before, however, this can be done, we must be in a position to estimate the number of particles that disappear from the chamber by subsidence and diffusion, while the process of aggregation is proceeding. As a first approximation, we can take the fall in weight of the dispersoid as a measure of this magnitude, and knowing what the weight would have been had there been no loss, multiply the observed numbers at any period by the ratio " maximum weight recorded/weight at given period." The numbers so obtained give an infinite cloud curve, that is to say, record the diminution in numbers that would be observed if the dimensions of the chamber were infinitely large, so that no loss by subsidence or diffusion could take place.

This calculation is obviously inapplicable to the early stages of the cloud, where the weight caught on the filter, in consequence of the growth of amicrons, gradually increases; but if the computation is made from the time when the maximum recorded weight begins to diminish, the fall in concentration can in many cases be taken as a measure of the number of particles lost by adhesion to the walls or by subsidence.

Again, it must be borne in mind that the fall in weight concentration is only a strict measure of the number deposited when the cloud consists either of even-sized particles or of particles which diffuse or fall at the same rate. In these clouds this is clearly not the case, but since the dispersoid is continually being mixed by stirring and convection, this factor is probably of secondary importance. A further uncertainty is caused by the determinations of weight and number referring to one cloud, whilst those of complexity were made on another.

Every precaution was taken to render the two clouds identical in character, and the conditions of dispersal were as nearly as possible the same.

Considering all these sources of error, the numbers in the following Table are as good as could reasonably be expected. The figures in col. 2 denote the particulate number at the times stated read from the curve (fig. 3), and those in col. 3 the corresponding weight concentrations taken from the micro-balance curve (fig. 4*a*). Col. 4 shows the particulate number for an infinite cloud obtained by combining the data in the two preceding columns. These numbers when plotted give a steeply descending curve, which turns and becomes nearly parallel to the time axis after a period of 3½ hours. The last column records the product of the particulate number and the degree of complexity, which should be constant if the changes in the infinite cloud are due solely to aggregation. In other words, it represents the number of particles per c.c. of an infinite cloud, if no aggregation takes place after the first half-hour.

Time.	Particles per c.c.	Weight (mgrm. per litre).	Number for infinite cloud.	Degree of complexity.	Corrected number.
h. m.					
0 30	$2\cdot86\times10^6$	0·0186	$2\cdot86\times10^6$	1·22	3·46
1 0	2·09	0·0175	2·22	1·41	3·13
1 30	1·62	0·0159	1·90	1·70	3·23
2 0	1·24	0·0144	1·60	2·00	3·20
2 30	0·95	0·0129	1·37	2·40	3·29
3 0	0·76	0·0115	1·23	2·79	3·43
3 30	0·63	0·0102	1·15	3·16	3·64
4 0	0·54	0·0091	1·10	3·49	3·84
4 30	0·48	0·0081	1·10	3·74	4·11
5 0	0·44	0·0075	1·09	—	—

It is hence clear from these results that the initial unstable period in zinc oxide clouds is caused, as far as the ultra-microns are concerned, by the process of aggregation, and this conclusion is almost certainly valid for all clouds of non-volatile oxides, and is most likely the determining factor also of the quick change in the other clouds examined.

Conclusion.

The results so far obtained can be summarised as follows:—

(1) Methods of determining the number, the average mass, and the size of the particles in highly diluted disperse systems in air have been worked out.

(2) A preliminary study of various aerosols has shown that in all the cases examined the smaller particles aggregate rapidly together, giving complexes of different structure, depending on the chemical nature of the substance dispersed. This process explains quantitatively the observed disappearance in the number of particles. Aggregation rapid when the cloud is first formed; ceases after a period varying from ½ to 4 hours.

(3) The striking fact has been discovered that all aerosols during the first stage of their life-history contain a relatively large proportion of matter, either in the molecular state or as amicroscopic particles.

At this stage of the investigation a comparison of the characteristics of aerosols and hydrosols would be out of place, but a parallel might be drawn between the process of aggregation and the quick coagulation of a gold colloid on the addition of a precipitating electrolyte. In the case of the hydrosol neutralisation of the charge brings about coagulation, and in the aerosol the particles are either electrically neutral or a proportion oppositely charged. Moreover, our particulate number curves are similar in form to the coagulation curves obtained by Zsigmondy (9) for gold and analysed mathematically by Smoluchowski (10).

It is hoped that a stricter examination of aerosols by the methods we have described will explain the cause of the process of aggregation, and throw further light on this analogy.

In conclusion, we wish to say how much we owe to Prof. F. G. Donnan, F.R.S., for his advice and encouragement. It was due entirely to him that the study of aerosols was first commenced, and we are indebted to him also for many ideas and suggestions during the later course of the investigation.

Our sincere thanks, too, are due to Prof. A. Smithells for much helpful criticism.

[*Note.*—Since writing these papers it has been found possible to apply Smoluchowski's theory of coagulation to the results we have obtained. With a slight modification Smoluchowski's equations reproduce the particulate number curves within the limits of experimental error.]

REFERENCES.

(1) Whytlaw-Gray, Speakman, Campbell. Pt. I of this paper.
(2) Millikan, 'The Electron,' p. 93.
(3) Einstein, 'Ann. d. Physik,' vol. 17, p. 549 (1905); Sutherland, 'Phil. Mag.,' vol. 9, p. 781 (1905).
(4) Millikan, 'The Electron' (1917).
(5) Wells and Gerke, 'J. Am. Chem. Soc.,' vol. 41, pp. 312-329 (1919).
(6) Steele and Grant, 'Roy. Soc. Proc.,' A, vol. 82, p. 580 (1909).
(7) Ramsay and Gray, 'Roy. Soc. Proc.,' A, vol. 86, p. 270 (1912).
(8) Knudsen, 'Ann. d. Physik,' vol. 50, p. 472 (1916); L. Hamburger, 'Koll. Zeitsch.,' vol. 23, p. 177 (1918).
(9) Zsigmondy, 'Zeit. f. Physik. Chem.,' vol. 92, p. 600 (1918).
(10) von Smoluchowski, 'Zeit. f. Physik. Chem.,' vol. 92, p. 129 (1918).

The Significance of the Electrode Potential.

By JAROSLAV HEYROVSKÝ, D.Sc., Ph.D.

(Communicated by Prof. F. G. Donnan, F.R.S. Received May 18, 1922.)

The potential of an electrode dipping into a solution of an electrolyte is usually explained by the Nernst osmotic theory, as resulting from the interaction of the so-called solution tension of the electrode material and the osmotic pressure of the ions in solution; but the actual mechanism which takes place at the electrode during the emission of ions, and the physical significance of the "solution tension" is left unexplained. In fact, this theory can hardly be taken as more than a formal analogy with osmotic phenomena, with which it shares thermodynamic treatment, and is consequently insufficient to explain any molecular mechanism of the process by which the electrode becomes charged.

The study of systems in equilibria, as ordinary reversible electrodes are, cannot give us an insight into such mechanism; this can, however, be obtained from the investigations of anomalous electrode potentials corresponding to processes tending to equilibrium. Thus the study of the passivity of magnesium led Kistiakowsky* to the view that the electrochemical process, which charges the electrode, is not

$$Mg \rightarrow Mg^{++} + 2\ominus$$

but

$$Mg + 2OH' \rightarrow Mg(OH)_2 + 2\ominus,$$

the symbol $\ominus$ denoting the negative electron.

Similarly, the potential of aluminium, as found by the writer,† corresponds to the process

$$Al + 3OH' \rightarrow Al(OH)_3 + 3\ominus.$$

Further, passive electrodes, on the whole, show sensitivity towards hydroxyl or hydrogen ions and oxidising or reducing agents, and are consequently regarded as kinds of oxygen electrodes working no doubt in a manner similar to that indicated above.

Thus the extraction of metallic ions from the surface of the electrode by means of hydroxyl ions, by which the electrode becomes charged seems to be a general primary process at any electrode in aqueous solutions.

Reichinstein‡ assumes, from the analogous behaviour of the galvanic as well as discharge-tube electrodes, that in both primarily oxygen ions are adsorbed to the metal, thus effecting the reaction

---

* 'Zeitsch. f. Phys. Chem.,' vol. 70, p. 206 (1910).

† 'Trans. Chem. Soc.,' vol. 117, p. 27 (1920).

‡ 'Zeitsch. Physikal. Chem.,' vol. 95, p. 457.

$$\mathrm{Me} + \mathrm{O}'' \rightarrow \mathrm{MeO} + 2\ominus.$$

He attempts to formulate this process thermodynamically, using several experimentally undeterminable constants without elucidating the nature of such a process or giving an explanation for the adsorption.

However, a deeper insight into the nature of electrochemical forces displayed at the electrodes can be obtained by introducing the physical conception of the electron affinity,* an idea which expresses in terms of the ionisation potentials the energy with which the valency electrons are held by the atoms.

The Origin of the Electrode Potential.

Consider an electrode consisting of a rod of metal, Me, dipping into an aqueous solution of its hydroxide, MeOH. At the surface of the metal, just as anywhere in the solution, the water molecules splitting into their constituent ions and rejoining, keep up a dynamic equilibrium, so that the ionic product $C_{H^\cdot} \times C_{OH'}$ is constant.

When a hydroxyl ion encounters a metallic atom at the surface of the metal, it may ionise the atom and join the metallic ion, thus forming a molecule, MeOH, which passes into the solution, whereas the negative electron is left at the metallic electrode. Such a process is

$$\mathrm{Me} + \mathrm{OH}' \rightarrow \mathrm{MeOH} + \ominus.$$

However, as the charging proceeds in this manner, the cations in the proximity of the electrode are attracted and deposited at the negatively charged electrode, and a constant potential is reached between the negatively charged electrode and the solution, at which the same number of atoms is continuously being extracted from the metal by hydroxyl ions as is deposited by the electrostatic attraction at the electrode. Let us express the work done in this process, regarding it as a reversible extraction of metallic atoms and subsequent deposition of the metallic ions at the electrode, thus:—

Vapourise one gramme-atom of the metal under its own vapour tension, p, which requires work, RT, and imagine this gramme-atom of the vapourised metal be brought into an ionisation chamber under the pressure, p. Further, bring into this ionisation chamber one gramme-molecule of water vapour from the solution surrounding the metallic electrode and expand it to the pressure, p, which requires work

$$\mathrm{RT} + \mathrm{RT}\log p/\mathrm{P}_{\mathrm{H_2O}}.$$

---

* Frank and Hertz, 'Phys. Zeitsch.,' vol. 20, p. 132 (1919).

Now ionise the metallic vapour, for which process an expenditure of work equal to $RT + I . F$ is necessary, where I denotes the ionisation potential and F Faraday's constant.

Also ionise the gramme-molecule of water vapour into $H^{\cdot}$ and OH' ions, the work necessary to effect this being $K_{HOH} . F$. We thus have in the ionisation chamber ions Me^+, $H^{\cdot}$ and OH', besides free electrons.

Allow the Me^+ ions to combine with the hydroxyl ions to form a gramme-molecule, MeOH; hereby energy $M . F + RT$ is gained. Now place the free electrons from the ionisation chamber into the metallic electrode by condensing them into the (let us say) platinum walls of the ionisation chamber, and by conducting them through a wire to the electrode; further, transfer the hydrions, as well as the molecules MeOH, from the ionisation chamber into the solution surrounding the metallic electrode.

If the concentration of the hydrions in this solution is $C_{H^{\cdot}}$ and that of the molecules MeOH is C_{MeOH}, the latter transfer involves work $RT \log C_{H^{\cdot}}/p + RT \log C_{MeOH}/p$, whereas to replace the electrons into the metal requires work $-\Phi . F + \phi - RT \log p$, where Φ denotes the energy of evaporation of the "electron gas" from platinum to unit pressure, and ϕ the Peltier's contact effect between the platinum and the metal; this, however, is very small, and may be neglected.

Before bringing the particles MeOH and $H^{\cdot}$ into the solution, they must be allowed to become hydrated; this involves energy changes

$$-H_{MeOH} . F - H_{H^{\cdot}} . F \text{ resp.}$$

In the solution, of course, the molecules MeOH will react with the hydrions according to the basicity of the hydroxide to form water and Me^+ ions, in the usual way:

$$MeOH + H^{\cdot} \rightarrow Me^+ + H_2O.$$

During this processs, which is going on at equilibrium concentration, no work is done.

The last step is to bring one gram-cation of Me^+ ions from the solution back to the metallic surface, which is done in transferring the ions through the potential difference π, which exists between the solution and the electrode. The work done here is $-\pi . F$.

Taking the work done by the system as positive, we obtain the total energy change in this reversible process as follows:—

$$nRT + RT \log p/P_{H_2O} + I . F + K_{HOH} . F - M . F + RT \log C_{H^{\cdot}}/p$$
$$+ RT \log C_{MeOH}/p - \Phi . F - RT \log p - H_{MeOH} . F - H_{H^{\cdot}} . F - \pi . F = 0.$$

To simplify this, express the work necessary to ionise a gram-molecule of water vapour, $K_{HOH} . F$, under the pressure p as the sum of:

(1) The work done in condensing the vapour from p to

$$P_{H_2O} = RT \log P_{H_2O}/p.$$

(2) Condensation to liquid water, $-RT$ and ionisation into $H^{\cdot}$ and OH' ions at equilibrium (work = 0).

(3) Transfer of dehydrated ions to the pressure p, which is

$$H_{H^{\cdot}} \,.\, F + H_{OH} \,.\, F + RT \log p/C_{H^{\cdot}} + RT \log p/C_{OH'}.$$

Substituting the sum of (1), (2), (3), for $K_{HOH} \cdot F$ we obtain

$$\pi = K - RT/F \log p + I - M + RT/F \log C_{MeOH}/C_{OH'} - H_{MeOH},$$

putting all the terms, which do not vary with the nature of the metal, into the constant K.

Further, applying the law of mass-action to the basic splitting of the hydroxide MeOH, we have $C_{OH'}/C_{MeOH} = K_B/C_{Me^{\cdot}}$, where K_B is the basic constant.

Considering that the energy of hydration, H_{MeOH}, which represents the energy change from vacuum to the aqueous solution, can vary but little for the different hydroxide* molecules, we obtain as the final result the formula (1):

$$\pi = K - RT/F \log p + I - M - RT/F \log K + RT/F \log C_{Me^{\cdot}}. \qquad (1)$$

This formula (1) includes all factors which influence the potential existing between the aqueous solution and the metallic electrode.

This potential is known to become less negative (more positive) if the vapour tension p of the metal be decreased (*e.g.*, by the addition of some solute); further, in general, all metals with a small ionisation potential I (*e.g.*, the alkali metals) have a very negative electrolytic potential; from these again those which form very compact molecules, *i.e.*, have a large value of M (*e.g.*, lithium), show the most negative value of potential; it is, further, familiar that the constant K_B of basic splitting is connected with the electrolytic potential inasmuch as a large negative value of the potential means strong basigenic properties.

Whilst the last term expresses the dependence of the potential on the concentration of ions as in the Nernst's formula, all the other terms summarised denote what is called in Nernst's osmotic theory the "solution tension," showing how complicated is the physical interpretation of this expression. From this we can understand why the electrolytic potentials are no simple function in the periodic system, although the ionisation potentials show great regularities in this respect.

---

* Le Blanc ('Abhandl. d. Bunsen Ges.,' No. 3, p. 72) finds indeed that hydration slackens the deposition of the metal, *i.e.*, makes the potential more negative.

The Order of Basicities and Beketoff's Principle.

In inorganic chemistry, especially in the theory of analysis, it is of the greatest importance to define in some manner the basigenic character of metals and to deduce the basic strength of a hydroxide from some physical properties. The heat of oxidation, the volume of oxides, the contraction during oxidation, or the heat of hydration of oxides, were proposed as a measure of basicities of oxides; however, none proved to be exact.

Abegg and Bodländer, in their theory of electro-affinity, suggested that the electrolytic potential might be a measure of the basic strength of the metal, but, as several exceptions show,* this idea cannot be regarded as quite correct. In general, Abegg expected the basigenity of a metal to depend on the position of the metal in the series of the electrolytic potentials, and, consequently, to vary according to the affinity for oxygen. However, from the researches of the Russian chemist, N. N. Beketoff,† on alkalies and alkaline earths, we can clearly see that alkalinity (or basigenity of the metal) has little to do with the affinity for oxygen, which manifests itself in the "stability of oxide."

Beketoff drew attention to the fact that the determining factor in the stability of a compound is the ratio of masses of combining atoms, in that a compound consisting of equally heavy equivalents of metal and metalloid is the most stable one; in other words, that its elements possess the greatest affinity for each other. Thus the oxides of heavy metals, rubidium, cæsium, barium—according to Abegg, the most "positive"—are easily reduced by metallic aluminium, its equivalent (= 9) being nearest to that of oxygen (= 8). For chlorides the stability is again different, the metals with equivalents nearest to that of chlorine being very stable and most difficult to reduce, etc. This influence of mass ratio upon the stability of compounds appears to be quite general, and may be called the *Beketoff's principle*; it seems, however, to apply to compounds now termed "ionogenic," possessing intramolecular ionisation, and being strong electrolytes. For such compounds we may express the chemical affinity, A, of the metal to the metalloid by means of electron-affinities as

$$A = E - I' + M.‡$$

---

* *E.g.*, the extreme value of the lithium potential, although lithium is the least positive of alkali metals; or the thallous potential, which is very near to that of hydrogen, in spite of the great basicity of the thallous hydroxide (*cf.* Heyrovský 'Trans. Chem. Soc.,' vol. 117, p. 1024, 1920).

† 'J. Russ. Chem. Soc.,' vol. 20, p. 525 (1888); see 'Journ. Chem. Soc. Abst.,' vol. 56, p. 332 (1889).

‡ In a similar way to Born's ('Verh. Deutsch. Phys. Ges.,' vol. 21, p. 19).

Here E denotes the electron-affinity of the metalloid atom to the electron, I′ the work necessary to ionise the metal, M the work necessary to ionise the molecule into its ions in the gaseous state. The quantity I′ differs from the ionisation potential of the single atoms by RT/F log p, where p denotes the tension of the metal.

Thus $$A = E+M-I+RT/F \log p.$$

Let us now return to formula (1) to substitute for the terms

$$I-M-RT/F \log p$$

the quantity E−A. Taking $C_{M_0^{\cdot}} = 1$, in which case π is the "electrolytic potential" (denoted by E.P.), we have

$$-RT/F \log K_B = \pi+A-E-K. \qquad (2)$$

Since K_B denotes the constant of the basic splitting of the hydroxide of the metal, this formula gives us a connection between the E.P. and the basigenity of the metal.

Consider now this relation for a series of metals, in order to obtain a simple approximate expression for their basigenities, which would assign a higher significance to the E.P., hitherto but of little theoretical significance.

At absolute zero, where no thermal agitations interfere in the stability of the molecule, the work M can be regarded as being purely electrostatic, viz., as the work necessary to remove the negatively charged particle (the anion) from the periphery of the metallic ion. Let this work be equal to the ionisation potential, I; then, according to the above equation, we obtain for absolute zero that

$$A_0 = E,$$

which is in agreement with experimental results* and theoretical deductions.†

For higher temperatures the thermal motion decreases the stability of an ionogenic molecule. Assume that this instability depends on the mass of the ions, so that the affinities A of metals to the same metalloid decrease, to be in agreement with the mentioned Beketoff's principle, with temperature in such a way that

$$A = E-kT \log m+k'T,$$

where m is the mass of the variable cation. Substituting now for A−E into the formula (2), we obtain finally for the basicity, B, of the hydroxide of a metal an approximate formula:

$$-B = -RT/F \log K_B = \pi-kT \log m+k'T-K. \qquad (3)$$

---

* Foote and Mohler, 'Journ. Amer. Chem. Soc.,' vol. 42, p. 1832.

† Born, 'Verh. Deutsch. Phys. Ges.,' vol. 21, p. 685.

In the following Table values of B are calculated from the known E.P. and the equivalent weights of the metals, the value of kT having been chosen empirically as 1·7 (at the room temperature).

The Order of Basigenities, B, of:

(1) The Monovalent Metals.

Metal.	E.P.	$B + K = 1{\cdot}7 \log_{10} m - \text{E.P.}$
Rb	−2·93	6·21
K	−2·92	5·63
Na	−2·71	5·03
Li	−3·02	4·45
Tl^{I}	−0·32	4·25
Ag	+0·80	2·66
Cu	+0·52	2·55
Au^{I}	+1·5	2·4

(2) The Divalent Metals.

Metal.	E.P.	$B + K' = 1{\cdot}7/2 \log_{10} m - \text{E.P.}$
Ca	−2·5	5·3
Pb^{II}	−0·12	4·05
Mg	−1·7	4·0
Mn^{II}	−1·0	3·96
Cd	−0·40	3·89
Zn	−0·76	3·84
Sn^{II}	−0·10	3·63
Fe^{II}	−0·43	3·40
Hg^{II}	+0·86	3·05
Cu^{II}	+0·34	2·72

The E.P. values are taken from the 'Abhandl. d. Bunsen Ges.,' No. 5 (1911), No. 7 (1914), those of the alkali metals from G. N. Lewis ('J. Amer. Chem. Soc.,' vol. 37, p. 1898).

It will be observed that the order of basigenitics comes out right, although the order of the E.P. is entirely different.

According to formula (3), the basicity, B, increases with temperature more rapidly for heavy metals than for the light ones; in fact, the hydroxides of lighter metals of such as Li, Be, Mg, Ca, do not show increase in solubility with temperature, whereas those of the heavy metals as Ba, Sr, K, are much more soluble at 100° C. than at 0° C.

Thus it can be taken as certain, from the above considerations, and from experience, *that the hydroxide of a metal is the more basic the more negative is the E.P. of the metal, and the heavier is its equivalent weight.*

The Mechanism of some Electrode Potentials.

The Process in the Cell.—Electrode processes taking place in short-circuited cells, such as Daniell's, are easily understood, when we imagine the stream of electrons as passing from the more negative zinc electrode to the less negatively charged copper electrode. Then the dynamic equilibrium at the zinc surface will be disturbed, and consequently the rate of Zn^{++} cations deposition will be slackened; as, however, the charging process proceeds with unaltered velocity, prevailing dissolution of the metal results. At the copper electrode, on the other hand, the negative charge is increased owing to the accession of electrons from zinc, and hence more Cu cations are attracted than at the equilibrium potential; consequently more copper is deposited than passes into solution.

Exactly the same happens if electrons are added or removed from electrodes externally, *i.e.*, by cathodic or anodic polarisation.

Concentration Cells.—The electrode dipping in the dilute solution of its cations has less chance to deposit cations than the electrode with the concentrated solution, whilst the negative charging of the metal, due to the ions of water, proceeds on either electrode with equal velocity; consequently, the dilute electrode maintains a more negative potential than the more concentrated one.

Electro-deposition.—The mechanism of electro-deposition of nobler metals from their salt solutions by immersion of a baser metal, for instance, the deposition of copper on metallic iron, can be explained as follows:—On any metal—as it always gets negatively charged—any of the cations present in solution are being deposited. However, the atoms deposited at the surface of the electrode are instantly acted upon by water molecules, which action brings the deposited ions back into the solution, preferentially those of very positive metals, according to the free energies involved in this interaction. The less positive then remain.

Over-voltage and the Hydrogen Potential.—From this point of view also the phenomenon of hydrogen over-voltage and other over-voltages might be elucidated. Since any cations existing in the surface layer are continually being deposited at the electrode, hydrogen ions also must be readily discharged from acid solutions as hydrogen atoms. However, the extraction of H atoms from the electrode to the solution proceeds with great ease, the energy of the process

$$H + OH' \rightarrow HOH + \ominus$$

being larger than in the case of the most positive metals; this corresponds to a very highly negative potential of a monatomic hydrogen electrode.

Consequently, we cannot get a steady deposition of hydrogen atoms—if their union to molecules proceeds comparatively slowly—unless a much more negative potential than the hydrogen zero is attained. The matter is, however, quite different if the surface of the electrode material is capable of at least a partial absorption of the atomic hydrogen. Then, owing to this activity—which we may call affinity of the electrode for hydrogen atoms—some of the deposited atoms of hydrogen are being drawn inside the surface, so that they can unite before they are extracted back to the solution and a steady evolution of the gas follows. In fact, the potential of the standard hydrogen electrode, regarded as a potential of atomic hydrogen, corresponds to very few atoms formed by the dissociation of molecules, the actual potential at which hydrogen atoms would be stable at the electrode being about as negative as that of most "positive" metals.* This position of atomic hydrogen in the electro-potential series is by no means discordant with the above rule of basigenity, if we consider the small equivalent weight of hydrogen.

Metalloid Electrode Potentials.

So far, the action of water upon metals only was considered as giving rise to the negative charge on the metal and to metallic ions in the solution. Such a union, however, between the hydroxyl ions of the solvent and atoms of the electrode is less probable in the case of electrodes consisting of non-metals, like chlorine, and is quite out of question for fluorine. Here, however, the other constituent of water, that is the hydrion, is apt to join the fluorine atom. The process in which fluorine atoms part from the electrode, leaving it charged, will be: firstly, association of fluorine atoms with the free electrons of the metallic electrode to negatively charged particles:—

$$\ominus + F \rightarrow F' \text{ (in the metal).}$$

Then the action of water at the surface of the electrode:

$$F' + HOH \rightarrow FH + OH'$$

and finally the dissociation:

$$FH \rightarrow F' + H^{\cdot}$$

the total process being:

$$\ominus + F \rightarrow F' \text{ (in the solution).}$$

Such an electrode can thus acquire only a positive charge, provided that the metalloid possesses sufficient negative electron affinity in order to furnish the requisite negatively charged ions in the metal. For a chlorine electrode, however, both processes, shortly written as

$$Cl + OH' \rightarrow ClOH + \ominus$$

and

$$Cl + H^{+} \rightarrow HCl + \oplus$$

---

* *Cf.* Benett and Thompson, 'Zeitsch. f. Electrochem.,' vol. 22, p. 233 (1916).

are possible. The resulting potential of the chlorine electrode might be either positive or negative, according to ionic concentrations. The other metalloids have smaller negative electron affinities, so that the prevailing process will be simply such as is the case of metals, for instance,

$$\mathrm{I} + \mathrm{OH}' \rightarrow \mathrm{IOH} + \ominus,$$

charging the electrode negatively.

It might be thought—this being the electrode process—that iodine would not be liberated at the anode, iodine atoms being rather sent into solution. However, if IOH enters the solution of an iodide, it at once reacts to form molecular iodine which deposits at the electrode; the total charge at this electrode (anode), corresponding to the passage of one Faraday, is then an increase of one equivalent of iodine, as is actually observed.

The different oxidation stages and uncertainties about electron affinities concerning metalloids render any quantitative calculations here impossible.

Electrode Processes in Non-aqueous Solvents.

The potential of an electrode dipping into a non-aqueous solvent can be regarded as being formed similarly to the manner in which aqueous electrodes become charged, for instance, in alcohols as being

$$\begin{array}{lc} & \mathrm{Me} + \mathrm{ROH} \rightarrow \mathrm{MeOH} + \mathrm{R}^{+} + \ominus \\ & \mathrm{MeOH} \rightarrow \mathrm{Me}^{\cdot} + \mathrm{OH}' \\ & \underline{\mathrm{R}^{+} + \mathrm{OH}' \rightarrow \mathrm{ROH}} \\ \text{total:} & \mathrm{Me} \rightarrow \mathrm{Me}^{+} + \ominus, \end{array}$$

or, in an amine, the MH_2' ion, taking the *rôle* of the hydroxyl ion.

The electrolytic potential in such solvents becomes, then, according to the formula (1),

$$\mathrm{E.P.} = -\mathrm{RT}/\mathrm{F} \log p + \mathrm{I} - \mathrm{M}' - \mathrm{RT}/\mathrm{F} \log \mathrm{K}_{\mathrm{MeR}} - \mathrm{H}_{\mathrm{MeR}} + \mathrm{K}',$$

which differs from the aqueous electrolytic potential in the term M', depending on the structure of the molecule, as well as in all following terms, which also depend on the dissociating power of the solvent.

This formula shows clearly that the electrolytic potential in non-aqueous solvents must differ from those in aqueous solutions; even their order might be different if the solvent give rise to ions other than hydroxyls. In some non-conducting and non-ionisable solvents there might be even no charge formed on the electrode. The experimental observations on potentials in non-aqueous solvents are, however, not sufficiently accurate to allow any quantitative interpretation.

The Absolute Electrode Potentials.

The charge acquired, according to the explanation put forward above, by any metal in aqueous solution always should be negative. Thus even electrodes of noble metals, such as silver or mercury, acquire a negative charge, owing to the action of water molecules. In the case of gold or platinum the value of the E.P. is uncertain, as these electrodes, which are irreversible with respect to their ions in solution, are rather apt to act as gaseous electrodes.* The least negative (or most positive) electrolytic potential exactly reversible is that of silver in a very concentrated solution of its ions, which is about 0·8 volt more positive (or less negative) than the hydrogen zero potential. On the other hand, the fluorine electrode, as was pointed out above, must be already positively charged, and, as its potential is nearly by 2 volts more positive than the hydrogen zero, the absolute zero potential must be somewhere between that of silver and fluorine, and most probably between silver and chlorine.

This is conflicting with the view still widely accepted, that the change of sign, *i.e.*, the absolute zero potential occurs at −0·28 volt, a value which was derived from electro-capillary phenomena. There is, however, no evidence which would show that this change of sign actually does occur, as it will be seen from the following consideration: There are two methods which lead to coinciding values of the "absolute potentials," that is, the determination of the potential at which "surface" tension (correctly interfacial tension) of polarised mercury is maximal, and, further, the measurement of the steady potential acquired by a dropping mercury electrode. But the second way is really identical with the first method, only here the mercury gets charged on spreading its surface in the electrolyte instead of being polarised externally; it follows necessarily from the thermodynamic principle of Le Chatelier-Braun, that if we form a mercury surface by applying external pressure upon the mercury, and if an electric charge is hereby acquired, the electric effect must be such as to increase the interfacial tension between the mercury and the solution, in order to oppose the growth of the surface of the mercury. Consequently, every sufficiently quickly dropping mercury electrode must automatically charge itself to the potential at which its interfacial tension in the solution is a maximum. This, of course, must coincide with the potential of maximal interfacial tension, to which in the same solution stationary mercury must be polarised, as it is well known from experience.†

The mercury at the maximum of interfacial tension might be charged to any potential, and indeed there are large differences of potential observed in cases when mercury drops into different solutions, although they should be all at absolute zero according to the Lippmann-Helmholtz theory. This erroneous point of view is due to disregarding the most important fact, that the interfacial tension is measured and not the surface tension of mercury.

---

* Their colloidal particles are charged negatively.

† The same applies to Christiansen's "moving drop" phenomena ('Ann. d. Phys.,' vol. 12, p. 1072 (1903)).

For the interfacial tension ($\gamma_{1,2}$) between mercury and the solution the following relation holds:—

$$\gamma_{1,2} = \gamma_1 + \gamma_2 - 2A^* \qquad (4)$$

where γ_1 is the surface tension of mercury and γ_2 the surface tension of the aqueous layer, A denoting adhesion between mercury and solution. All these terms change with the polarisation (π), and consequently, instead of the Lippmann-Helmholtz equation:

$$\gamma_1 = -\tfrac{1}{2}C\pi^2 + \kappa,$$

we shall have, in the simplest case, assuming that γ_2 changes with polarisation similarly as γ_1,

$$\gamma_{1,2} = -C\pi^2 + \int f(\pi)\,d\pi + \kappa \qquad (5)$$

where $f(\pi) = dA/d\pi$ represents the change of adhesion with polarisation.

It is also evident from the formula (4) that substances affecting the surface tension of water must produce changes of the interfacial tension, and hence give different values for the Nernst-Ostwald zero potential.

In recent papers published on the electro-capillarity of mercury† this problem is being attacked from the view-point of the relationship as given in (4), although already, in 1903, Kučera‡ thoroughly discussed the important *rôle* of adhesion and the surface tension of solution displayed in these phenomena.

It is, however, still assumed that near the maximum of the electro-capillary parabola the sign on mercury changes. This assumption is based upon the concordant values of zero potentials and the nearly ideal shape of the parabolas, if substances influencing surface tension are absent. However, relation (5) shows that in such cases we might expect parabolas giving coincident values of potentials corresponding to the maximum if $f(\pi)$ is a constant or a linear function of π. Then $d\gamma/d\pi = k'\pi + k$, which gives for the maximum condition of interfacial tension equation $\pi = -k/k'$, with a potential different from zero.

It may be added that, although much attention was paid at one time to the exact determination of the absolute zero, which was believed to be at −0·56 volt, from the potential of a standard calomel electrode, yet this idea has proved theoretically altogether fruitless.

---

* Volkmann, 'Wied. Ann.,' vol. 16, p. 321 (1882).

† Krüger and Krumreich, 'Zeitsch. f. Electrochem.,' vol. 19, p. 617 (1913); Gouy, 'Ann. de Phys.,' vol. 9, p. 129 (1917); Frumkin, 'Phil. Mag.,' vol. 40, p. 363 (1920).

‡ Ann. der Phys.,' vol. 11, p. 529.

My thanks are due to Prof. B. Brauner for directing my attention to the works of Beketoff.

Summary.

Evidence is adduced that the process by which a metallic electrode, dipping into an aqueous solution, becomes charged is

$$\mathrm{Me + OH' \rightarrow MeOH + \ominus}.$$

The value of the electrode potential is deduced by means of the electronic conception of chemical combination, and a formula is obtained which expresses the electrolytic potential in terms of the ionisation potential and the basigenity of the metal.

The influence of the equivalent weight of a metal upon the basicity of its oxide is discussed.

Various electrode processes are explained from the above view-point.

The "absolute zero potential" value derived from the electrocapillary phenomena is shown to be illusory, and the probable potential, at which the change of sign takes place, is indicated.

The Effect of Long Grinding on Quartz (Silver Sand).

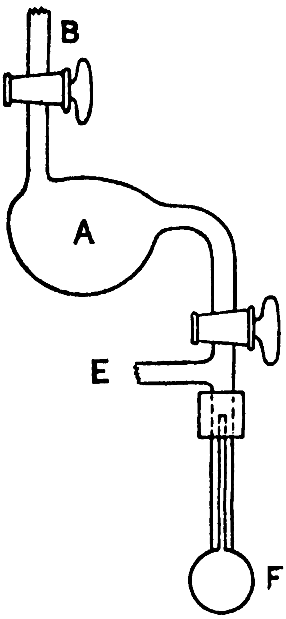
By Rames C. Ray, M.Sc.

(Communicated by Dr. M. W. Travers, F.R.S. Received August 11, 1922.)

It has been shown* from determinations of the heats of solution that when silver sand is subjected to grinding for a long time in a mechanically operated agate mortar, it partially loses its crystalline structure and is converted into the vitreous state. The value of the molecular heat of solution of quartz (silver sand), after grinding for 15 hours, was found to be 2·17 kgrm.-calories higher than that obtained for material which had not been subject to grinding. The difference between the molecular heats of solution of quartz and vitreous silica is 6·95 kgrm.-calories, so that 31·2 per cent. of the material is converted into the vitreous state by 15 hours' grinding. In the course of the discussion on the paper, it was suggested that, as there was a considerable difference between the densities of quartz and silica glass, the measurement of the densities of the ground and unground materials ought to show a difference, if any considerable change from the crystalline to the vitreous state really took place. The following experiments were undertaken in order to settle this point.

---

* 'Roy. Soc. Proc.,' A, vol. 101, p. 509 (1922).



In the preliminary experiments, it was found that consistent results could not be obtained in the case of the ground silver sand, when water was used as the liquid for determining the volume of the material. Considerable difficulty was also experienced in removing the adhering air when attempt was made to determine the density in the usual way with a specific gravity bottle. The following method, which was finally adopted for determination of the densities, gave fairly good results: A small bulb of about 10 c.c. capacity was blown at the end of a short capillary tube, the internal diameter of which was about 0·5 mm. The bulb was steamed out for nearly 20 minutes, and then heated at 300° for 24 hours so as to eliminate contraction of the glass. The capacity of the bulb was determined by filling it up with water. The liquid used for determining the volume of the materials was carbon tetrachloride, which had been re-distilled after digesting with calcium chloride for a long time. All the densities were determined at the air temperature 16°–17·6° C. The following values were obtained for the density of carbon tetrachloride:—

$$\left.\begin{array}{ll} (1) \ \ldots\ldots\ldots\ldots\ldots\ldots & 1\cdot597 \\ (2) \ \ldots\ldots\ldots\ldots\ldots\ldots & 1\cdot601 \\ (3) \ \ldots\ldots\ldots\ldots\ldots\ldots & 1\cdot599 \end{array}\right\} \text{Mean, } 1\cdot599$$

For determining the densities, the same sample of materials which had been employed for the measurements of the heats of solution (*loc. cit.*) was again dried at 150° C. The apparatus employed was a modified form of that described by Wade and Merriman,* and is shown in the accompanying diagram. Carbon tetrachloride, contained in the vessel A, was boiled under reduced pressure, and, after cooling, allowed to run into the evacuated bulb F, which contained a weighed quantity of the material. The tubes B and E served for connection with a Töpler pump. The results obtained are tabulated below, densities being referred to water at 4° C.:—

---

* 'Jour. Chem. Soc.,' vol. 101, p. 2430 (1912).

Substance.	Density.	Mean.
Silver sand (unground)	2·638 2·640 2·637	2·638
Silver sand (ground for 15 hours)	2·530 2·526 2·528	2·528
Silica glass (crushed)	2·208 2·206 2·210	2·208

The above results clearly indicate that a considerable amount of the crystalline material is converted into the vitreous condition on grinding. The difference between the densities of silver sand (unground) and silica glass is 0·43, and of silver sand unground and silver sand which has been ground for 15 hours is 0·11. Thus, from the consideration of the change of density from the unground to the ground silver sand, 25·7 per cent. of the material is converted into the vitreous state. This value is of the same order of magnitude as obtained from the determinations of the molecular heats of solution in hydrofluoric acid. As the ground material must be under a strain, a closer agreement between the two values cannot possibly be expected.

In this connection, it is interesting to note that Brown, Cook and Warner,* who have studied the effect of grinding on the apparent density of lead oxides, have observed a similar diminution in the density of crystalline lead oxide by grinding. The decrease of density is, however, not shown by amorphous lead oxide.

Summary.

It is found that, when quartz (silver sand) is ground for a long time, the density of the ground substance is lower than the one which has not been subjected to grinding. The fall of density shows that as much as 25·7 per cent. of the material is converted from the crystalline to the vitreous condition. This value agrees fairly with that arrived at from the determination of the molecular heats of solution.

Finally, I wish to express my thanks to Dr. M. W. Travers, F.R.S., and Prof. F. G. Donnan, F.R.S., for the interest they have taken in this work.

---

* 'Jour. Physical Chemistry,' vol. 26, p. 480 (1922).

BAKERIAN LECTURE: *The Distortion of an Aluminium Crystal during a Tensile Test.*

By G. I. TAYLOR, F.R.S., and C. F. ELAM.

(Received December 16, 1922,—Lecture delivered February 22, 1923.)

The work described in the following pages was inspired by a paper in which Prof. Carpenter* and Miss Elam described the result of applying tensile tests to specimens of aluminium which had been treated in such a way that they appeared to turn into single crystals. The resulting distortions of the test pieces were very remarkable and clearly suggested that the crystal axes were not orientated in the same direction in different specimens. The uniformity of the distortion in different parts of the same specimen made it seem likely that it would be a straightforward, though possibly laborious, matter to determine the relationship between the orientation of the axes and the distortion produced in a tensile test. And it seemed possible that by examining a number of specimens some general results might be obtained about the forces necessary to produce distortions of this type.

On discussing the matter with Prof. Carpenter and Miss Elam it was found that it would not be possible to determine the distortion from the measurements they had already made. Moreover, no measurements of the orientation of the crystal axes had been made, though Sir W. Bragg had made a few observations indicating that the material retained its crystalline character after it had been distorted. Under these circumstances, it was decided to carry out a test, making all the necessary measurements at various stages during the extension of a specimen.

Before describing the test, however, it is necessary to refer to the work of previous experimenters on the subject.

Previous Work.†

When a metal is strained beyond the elastic limit a microscopic examination of the surfaces of crystals in it frequently shows the existence of lines known as slip bands. These bands have been shown to consist of small steps, and the conclusion is naturally drawn that there are planes inside the crystal, presumably crystal planes, on which slipping takes place. The bands would then mark the intersection of the face of the specimen with these crystal planes.

---

* These 'Proceedings,' vol. 100 (No. A 704, December, 1921).

† See Note added at end of Paper.

Up to the present, however, the evidence on slipping is purely qualitative. It has not been shown that the deformation of a metallic crystal when the material is strained is such as could be produced by slipping, nor has the relationship between the crystal axes and the slip planes been determined.

A number of experiments have been made on the direction of the crystal axes in drawn wires, and certain conclusions have been reached by Polanyi* in regard to the orientation of the crystal axes with respect to the axis of the wire. References are given in Polanyi's paper.

General Description of Test.

Before the test contemplated could be carried out, it was necessary to enlist the help of an expert in crystal analysis by X-rays. Fortunately, Dr. Alex Müller took up the work and succeeded in devising a satisfactory method of determining the orientation of the crystal axes. This method and some of his results will be described elsewhere.

The results of his and Miss Elam's X-ray analysis of the specimen with which the present work is concerned are given in Table V.

Most of the specimens with which Prof. Carpenter and Miss Elam's experiments had been carried out were flat strips, about $\frac{1}{8}$ inch thick × 1 inch broad. These strips were unsuitable for the present purpose, partly because of the difficulty of making accurate measurements on the narrow faces and partly because they were so broad in comparison with their length that it was not possible to make sure of getting an evenly stretched parallel piece in the middle.

Specimens could be produced with circular, but not with square sections. On the other hand, there seemed to be no very simple way of making on a round specimen† the measurements which are necessary for calculating its distortion. For this reason, therefore, a round specimen was machined down till its section was square. Its dimensions were then approximately $1 \cdot 0 \times 1 \cdot 0 \times 20 \cdot 0$ cm. Each face was marked by a scratch parallel to the length of the specimen or axis, as it will be called, and by cross scratches. The appearance of the specimens so marked is shown in fig. 1. The faces were numbered 1, 2, 3, 4, so that when the specimen was placed upright in the testing machine, the faces appeared in this order when the observer moved round the machine in an anti-clockwise direction. Fig. 1 represents the specimen lying with its top end to the right.

At each successive stage of the test the extension between each pair of cross marks was measured on each face. The ratio of the length at any stage to the initial length will be denoted by the symbol ϵ. At the same time the angles between these cross scratches and the longitudinal scratch were measured in each case. These will be denoted by β on face 1 and γ on face 4.

---

* Polanyi, 'Die Naturwissenschaften,' April, 1922.

† This difficulty has now been overcome.

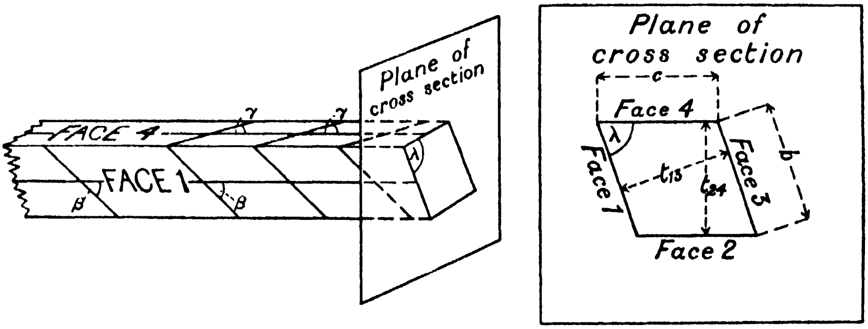


FIG. 1.—Scheme for marking and measuring specimen.

The thickness of the specimen between pairs of opposite faces, t_{13} and t_{24}, and the angles, λ, between neighbouring faces were also measured. These measurements are sufficient to determine the nature of the distortion. The scheme is illustrated in fig. 1.

Methods of Measurement.

The measurements of extension were made with a reading microscope, and the measurements of thickness between opposite faces with a micrometer. The angles between the cross lines and the axis of the specimen were measured with a crystallographer's microscope, with rotating eyepiece, containing cross wires, which was kindly lent us by Mr. A. Hutchinson, F.R.S. The angles between the faces were measured by sticking small pieces of cover-slip glass to them with gum, setting the specimen upright in a goniometer, and observing the reflection in the glass of a distant source of light.

Degree of Uniformity of Measurements.

If the distortion of the specimen had been uniform throughout its entire volume, and if the section had been accurately rectangular to begin with, the angles of the cross scratches measured on pairs of opposite faces would have been the same, the extensions measured on all four faces would have been the same, and the angles between any pair of faces would have been the same as that between the opposite pair.

It turned out that near the ends of the specimen, where it was held in the grips of the testing machine, the measurements were not quite the same as those near the middle, but that the central portion was nearly uniformly strained. The degree of uniformity can be judged from the figures given in Table I, which represent the ratio of the extended to the initial lengths in each of the five compartments into which the faces were divided by the six cross lines. The ratio will be represented by the symbol ϵ, $\epsilon-1$ being the extension.

In Table I the figures in each vertical column represent the values of ϵ for one of the eight stages of the test at which the specimen was measured. These stages correspond roughly with 5, 10, 15, 20, 30, 40, 60 per cent. extension, and the last column represents the extension after the specimen had broken at 78 per cent. extension.

Similar tables were prepared for the other measurements, but it seems unnecessary to print them in detail as they show a similar type of uniformity among the measurements taken in the central part of the specimen.

In fig. 2 is shown a photograph of part of the specimen after it had broken. It will be seen that its appearance suggests that the distortion is due to a uniform strain at all points, except those which lie in a small region close to the point where the breakage occurred.

Fig. 2.

Dimensions of Mean Parallelepiped.

On looking at Table I it will be seen that there is a high degree of uniformity among the measurements of extension in the three middle compartments on each face, especially in the earlier stages of the test.

Similar conclusions were reached by inspection of the other measurements ;

Table I.—Ratio of Extended to Initial Length in each of the Five Compartments into which each Face was Divided.

Approximate extension	5 per cent.	10 per cent.	15 per cent.	20 per cent.	30 per cent.	40 per cent.	60 per cent.	78 per cent.
Face 1	1·049	1·103	1·157	1·196	1·300	1·401	1·623	—
	1·047	1·108	1·160	1·199	1·309	1·414	1·645	1·834
	1·058	1·116	1·164	1·200	1·303	1·398	1·604	1·756
	1·052	1·109	1·160	1·199	1·300	1·394	1·595	1·739
	1·052	1·109	1·153	1·278*	1·297	1·390	1·582	1·692
Face 2	1·043	1·100	1·141	1·185	1·283	1·384	1·599	—
	1·048	1·106	1·160	1·199	1·300	1·408	1·628	1·805
	1·053	1·112	1·162	1·201	1·311	1·412	1·645	1·845
	1·051	1·107	1·161	1·198	1·295	1·393	1·649	1·720
	1·052	1·109	1·161	1·202	1·308	1·406	1·561	1·735
Face 3	1·046	1·099	1·153	1·191	1·296	1·392	1·618	1·828
	1·051	1·111	1·162	1·202	1·311	1·424	1·650	1·845
	1·059	1·112	1·163	1·201	1·304	1·401	1·613	1·760
	1·054	1·110	1·163	1·201	1·298	1·397	1·596	1·800
	1·056	1·111	1·150	1·198	1·300	1·390	1·581	1·611
Face 4	1·051	1·103	1·157	1·200	1·304	1·410	1·641	—
	1·050	1·111	1·169	1·203	1·316	1·420	1·648	1·830
	1·056	1·112	1·152	1·200	1·300	1·400	1·615	1·772
	1·054	1·109	1·159	1·198	1·299	1·394	1·594	1·720
	1·058	1·111	1·080	1·201	1·300	1·393	1·591	1·705
Means of middle three	1·053	1·110	1·161	1·200	1·304	1·404	1·623	1·785

* Probably an error in measurement.

accordingly, the mean has been taken in each case of the measurements made in each of the three central compartments. The mean extension, for instance, at any stage is the mean of the twelve figures obtained on all four faces. These are given at the bottom of Table I.

The mean angle of the cross line with the axis of the specimen is found by taking the mean of the eight measurements made on the four middle cross marks on two opposite faces.

In this way the dimensions and angles of a series of mean parallelepipeds have been drawn up, and these have been used in calculating the strain, or distortion, of the specimen at any stage. They are given in Table II.

It is known that when aluminium is worked the change in density is small, at any rate it is less than 1 per cent. The volumes of the strained figures should differ by less than 1 per cent. from the volumes of the unstrained figure. The ratio of the strained to unstrained volumes has been calculated from the figures in Table II. They are given at the bottom of Table II, and on inspecting them it will be seen that none of them differ from 1 by more than 1 per cent., except the figure in the last column but one corresponding to an extension of 62 per cent. This differs from 1 by nearly 8 per cent. It seems probable, therefore, that some mistake has been made in this case, and all the measurements for this stage have accordingly been rejected.

Table II.—Dimensions of Mean Parallelepiped.

Extension		0	1·053	1·110	1·161	1·200	1·304	1·404	1·623	1·785
Thickness in mm.	t_{13}	10·16	9·68	9·17	8·82	8·55	7·90	7·38	6·57	6·03
	t_{24}	10·35	10·33	10·31	10·28	10·26	10·12	10·00	10·17	9·29
Angle between Faces 1 and 4, λ°		90·6	88·8	87·2	86·7	84·9	81·8	79·2	73·1	70·8
Angle between cross marks and axis	Face 1, β°	90·0	88·5	87·3	86·7	85·7	84·2	83·0	80·2	78·0
	Face 4, γ°	90·0	90·3	89·7	89·3	88·5	85·2	82·1	73·2	65·5
Volumes		1·000	1·002	0·990	1·003	1·005	1·002	1·003	1·078	1·007

Analysis of Strain in Mean Parallelepiped.

It is obvious that the measurements given in Table II are sufficient to determine the strain completely. A uniform strain can be specified by giving the directions and magnitudes of the axes of the ellipsoid into which a sphere of unit radius in the unstrained material is transformed by the strain. This ellipsoid, known as the "strain ellipsoid," can be found from the measurements of Table II, but it is not at once obvious how the axes of the strain ellipsoid would be related to the crystal axes, though there would probably be some indirect connection between the two. On the other hand, there are other possible ways in which a uniform strain could be specified, and some of these may be more likely to throw light on the present problem than others.

For this reason it seems desirable to consider what types of strain can be conceived which would satisfy the conditions indicated by experiment, namely, that the material remains a crystal, so that all molecules are orientated in the same direction at any stage in the strain, and that the density is practically unchanged. In the first place, it seems clear that the relative displacements of neighbouring molecules cannot be the same as the relative displacements of particles in a similar position in the material in bulk. If they were the same, that is, if the strain were uniform even when portions of matter of molecular dimensions were examined, the material would remain crystalline when subjected to any uniform strain (it is true) but the crystal symmetry would be altered by the strain. A plane of molecules which lay in a crystal plane in the unstrained material would continue to do so in the strained material, and the angle between two such planes of molecules would alter continuously during any continuously varying strain. This state of affairs is mathematically conceivable, but it is contrary to all physical experience; moreover, measurements of the inclination of two different crystal planes during the present experiments showed that the angle between them remained practically constant during the whole course of the distortion. (See last column Table V.)

The fact that the material preserves its cubic crystal symmetry during the distortion, leads to the conception that the apparently uniform strain of the material in bulk must be made up of a large number of non-uniform strains or relative displacements between neighbouring molecules.

In order that crystal symmetry may be preserved, the displacement of any molecule relative to a neighbouring molecule must be at least as great as the distance between neighbouring molecules in the direction of its displacement. A small strain in the material in bulk must therefore be due in some way to the occurrence of a small number of these relatively large displacements, the greater part of the molecules preserving their relative positions unchanged.

When a molecule is displaced along any line into the position of a neighbouring molecule, the molecule which previously occupied that space must displace in its turn a third molecule and so on. Owing to the fact that the stress in a material under a uniform tension is constant along any line, it seems likely that all the molecules in a line would be displaced together. By making appropriate combinations of displacements of lines of molecules parallel to a given line in the crystal, the material in bulk might be conceived to assume a great variety of shapes, but they would all be characterised by the fact that lines of particles parallel to this direction would be unchanged in length during the distortion.

These considerations, though admittedly hypothetical, suggest that it would be more promising to analyse the strain, with a view to finding the directions in the material which are unchanged in length than to find the strain ellipsoid. In a uniform strain, the directions which remain unchanged in length lie on a quadric cone in the strained material. This cone evidently passes through the curved line of intersection of the strain ellipsoid and the unit sphere when these two are placed concentrically. The principal axes of this cone therefore coincide in direction with those of the strain ellipsoid.

The particles of this "unstretched cone" have evidently two positions corresponding with the unstrained and the strained material respectively. It is necessary to determine both of them.

Determination of the Cones of Unextended Directions.

The method adopted was to find first a series of corresponding positions in the strained and unstrained material respectively of planes of particles which pass through one edge of the specimen; actually the intersection of faces 1 and 4 was chosen. This was accomplished as follows :—

The cross-sections of the strained and unstrained specimen were set out on a piece of paper in such a way that the lines representing face 1 coincided in direction, while the points representing the edge of intersection of faces 1 and 4 coincided. The figure produced in this way is shown in fig. 3; OPQR is the cross-section of the unstretched specimen, while OSTU is that of the stretched specimen. Evidently, OR, OU, which are the traces of face 1, and OP, OS, which are the traces of face 4, are the traces of two pairs of corresponding planes; the diagonals OQ, OT are another pair. To find others, take any point V in QR, and let n represent the ratio RV : RQ. Divide UT so that UW $= n$UT. Then OV and OW are the traces of corresponding planes.

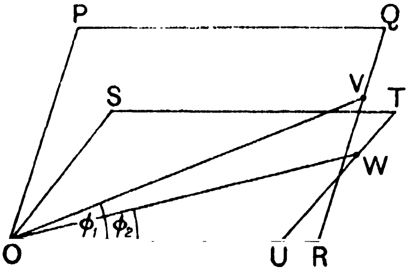


Fig. 3.

Let λ_1 and λ_2 be the angles between face 1 and face 4 before and after straining, namely, POR and SOU in fig. 3, let b_1, b_2 represent the breadth of the face 1 before and after straining, so that OR is b_1 and OU is b_2, and let c_1, c_2 represent the corresponding breadths of face 4. If ϕ_1 and ϕ_2 are the angles between corresponding planes in the unstrained and strained material and face 1 in each case, and if d_1, d_2 represent OV, OW, the breadths of the corresponding longitudinal sections, then

$$\left.\begin{aligned} d_1^2 &= b_1^2 + n^2 c_1^2 + 2 n b_1 c_1 \cos\lambda_1 \\ d_2^2 &= b_2^2 + n^2 c_2^2 + 2 n b_2 c_2 \cos\lambda_2 \\ \sin\phi_1 &= n c_1 \sin\lambda_1 / d_1 \\ \sin\phi_2 &= n c_2 \sin\lambda_2 / d_2 \end{aligned}\right\} . \qquad (1$$

Varying n from $-\infty$ to $+\infty$, all possible pairs of corresponding planes through the edge where faces 1 and 4 intersect are found.

The next step is to find the directions in corresponding planes which are unchanged in length by the strain. These are evidently the intersections of the cones of unstretched directions with this pair of corresponding planes.

At this stage it is necessary to make use of the observed angles between the cross lines on faces 1 and 4 and the axis of the specimen. If corresponding values of these are β_1, β_2 on face 1 and γ_1, γ_2 on face 4, it is possible to find, geometrically or analytically, the angles α_1, α_2 between the axis of the specimen and the lines of intersection of the planes containing the cross lines on faces 1 and 4 and any pair of corresponding planes through the edge where faces 1 and 4 intersect. The relation between α, β, γ, ϕ, and λ is shown in fig. 4.

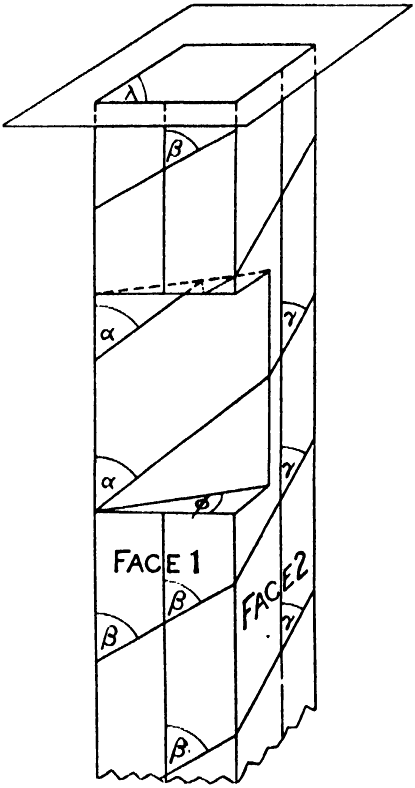


Fig. 4.—Diagram showing how α and ϕ are related to the faces of the specimen.

Next by rotation, about the line of intersection of faces 1 and 4, bring the two corresponding planes into coincidence. On these two coincident planes construct two corresponding parallelograms, ABCD, AEFG (see fig. 5), as follows:—

Draw two lines, BC, EF, parallel to AD (which is on the line of intersection of faces 1 and 4), and at distances d_1 and d_2 from it. At the point A draw two lines, AB, AE, at angles α_1 and α_2, to AD, so that BAD $= \alpha_1$, EAD $= \alpha_2$. On AD take any length AD in the unstrained plane. The length AG of the corresponding side of the strained rectangle is found by taking AG $= \epsilon$(AD), where ϵ is the ratio of the lengths of the specimen after and before stretching. ABCD and AEFG are the unstrained and strained shapes of a parallelogram of particles in the material.

To find the pairs of lines which are unstretched by the strain:—

Join BE and CF, fig. 5, and let them cut in H. Any straight line through H cuts the lines BC and EF in two points, K_1 and K_2 say, which are such that AK_1 AK_2 are corresponding lines of particles in the unstrained and strained figures respectively.

To find the positions of K_1 and K_2 which are such that $AK_1 = AK_2$, draw a line midway between BC and EF and parallel to them; let it cut a circle described on HA as diameter in I and J. Join HI, HJ, and let these lines cut BC, EF in L_1, L_2, and M_1, M_2. Evidently $AL_1 = AL_2$ and $AM_1 = AM_2$, so that these lines are in the required directions.

These operations can be performed analytically, and it is sometimes more convenient to do so. Determining d_1, d_2, ϕ_1, ϕ_2 from the measured quantities b_1, b_2, c_1, c_2, λ_1, λ_2 by means of equation (1), the angles α_1, α_2 are given by

$$\left.\begin{aligned} d_1 \cot \alpha_1 &= b_1 \cot \beta_1 + n c_1 \cot \gamma_1 \\ d_2 \cot \alpha_2 &= b_2 \cot \beta_2 + n c_2 \cot \gamma_2 \end{aligned}\right\} . \qquad (2)$$

If the angles L_1AD, M_1AD of the two unstretched lines of particles in the unstrained material be θ_1, θ_1', and if θ_2, θ_2' represent the corresponding angles L_2AG, M_2AG in the strained material, θ_1, θ_1' are the two values of θ_1 given by

$$\cot \theta_1 = \frac{\epsilon}{\epsilon^2 - 1} (\epsilon \cot \alpha_1 - f \cot \alpha_2) \pm \frac{1}{\epsilon^2 - 1} \sqrt{(\epsilon^2 - 1)(1 - f^2) + (\epsilon \cot \alpha_1 - f \cot \alpha_2)^2}, \qquad (3a)$$

and θ_2, θ_2' are the two values of θ_2 given by

$$\cot \theta_2 = \frac{1}{\epsilon f} \left\{ \frac{\epsilon}{\epsilon^2 - 1} (\epsilon \cot \alpha_1 - f \cot \alpha_2) \pm \frac{\epsilon^2}{\epsilon^2 - 1} \sqrt{(\epsilon^2 - 1)(1 - f^2) + (\epsilon \cot \alpha_1 - f \cot \alpha_2)^2} \right\}, \qquad (3b)$$

where f represents the ratio d_2/d_1.

Thus we can determine the spherical polar co-ordinates (θ_1, ϕ_1) and (θ_1', ϕ_1) of two directions in the unstrained material which after straining are unextended. Their new co-ordinates are then (θ_2, ϕ_2) and (θ_2', ϕ_2). By taking all possible values of n in formulæ (1) and (2) we can map out the two corresponding cones of unstretched directions.

Representation of Unstretched Cones.

The natural way to represent these cones is on a sphere, but for convenience it is necessary to use some flat representation of the sphere. The most convenient of these is the stereographic projection, which is the projection from a point on the surface of the sphere on to the plane through the centre parallel to the tangent plane at the point of projection.

The plane of projection is divided by the sphere into two parts. The inner part represents the hemisphere which lies on the opposite side of the plane of projection to the point of projection. As a rule, a convention is adopted which enables the whole sphere to be represented by the area inside the circle of intersection of the sphere and plane of projection. Assuming the plane to be horizontal, all points of the hemisphere above this are projected from the lowest point of the sphere, while all points of the lower hemisphere are projected from the highest point of the sphere. The representations of the two hemispheres then cover the same area, and they are distinguished by representing points of the upper hemisphere with a dot, and points of the lower hemisphere with a cross. In the diagrams which follow, curves on the upper hemisphere are represented by full lines through the dots, while curves on the lower hemisphere are represented by dotted lines through the crosses. With these conventions a great circle on the sphere is represented by a lens-shaped figure, consisting of two equal circular arcs, which pass through opposite ends of a diameter of the bounding circle of the projection, and lie on opposite sides of it. One of these arcs is a full line and the other is dotted. Several of them will be seen in figs. 7 and 8.

The usefulness of the stereographic projection is greatly increased if a figure known as a "stereographic net" is used. This consists of the representation of a series of meridian circles and parallels of latitude projected from a point on the equator, so that the two poles of the system lie on the bounding circle of the projection. Such a net is shown in fig. 6.

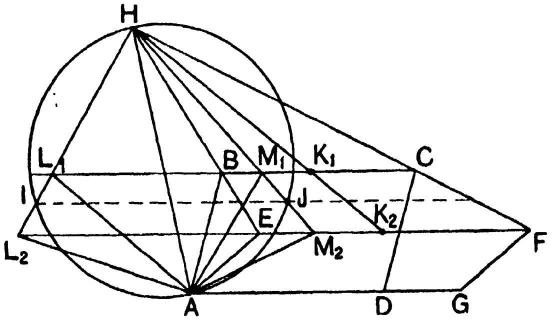


FIG. 5.—Construction for finding directions which are unchanged in length by strain.

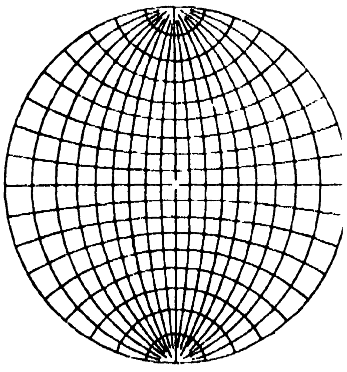


FIG. 6.—Stereographic Net.

To set out a point on a stereographic projection when its spherical polar co-ordinates (θ_1, ϕ_1) are given, let the pole of the spherical co-ordinates, *i.e.*, the point $\theta = 0$, be represented by the centre of the stereographic figure. Draw a circle on a piece of tracing paper equal to the bounding circle of the stereographic net, and mark a radial line on it to represent the meridian circle $\phi = 0$. Next draw another radial line at angle ϕ_1 with the line to represent the meridian $\phi = \phi_1$. Then lay the tracing paper over the stereographic net so that the line $\phi = \phi_1$ coincides with one of its two principal diameters, and mark off a point to represent the angle $\theta = \theta_1$, using a dot if θ_1 is less than $\frac{1}{2}\pi$ and a cross if it is greater. The points in figs. 7 and 8 are plotted in this way.

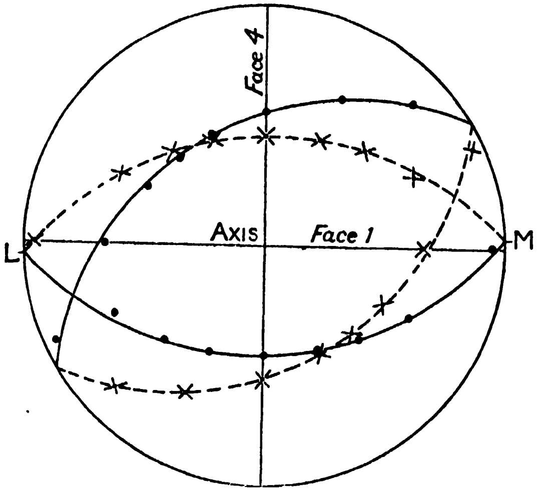


FIG. 7.—Stereographic projection of positions in *unstrained* material of cone of directions which are unstretched after the material has stretched by 30 per cent.

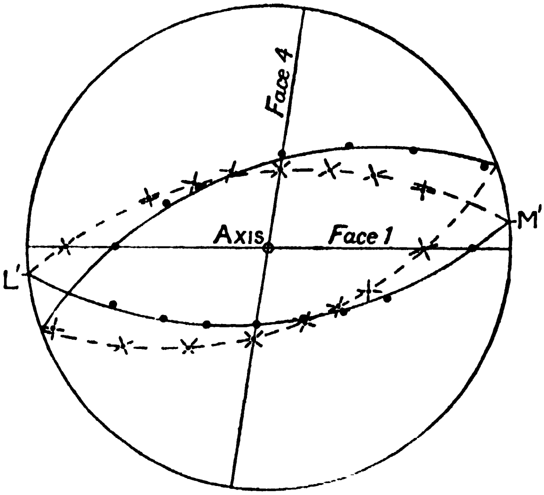


FIG. 8.—*Strained* position of cone shown in fig. 7.

Application of the Method in the present Case.

To analyse a strain by this method, two stereographic figures must be made, showing the undistorted and the distorted positions of the unstretched cone.

In the first instance the distortion produced by a 30-per-cent. extension of the specimen was analysed. Using the following values from Table II, $b_1 = t_{24}/\sin\lambda_1 = 1{\cdot}035$, $b_2 = t_{13}/\sin\lambda_2 = 1{\cdot}022$, $c_1 = 1{\cdot}016$, $c_2 = 0{\cdot}798$, $\lambda_1 = 90^\circ\ 35'$, $\lambda_2 = 81^\circ\ 50'$, $\beta_1 = 90^\circ$, $\beta_2 = 84^\circ\ 10'$, $\gamma_1 = 90^\circ$, $\gamma_2 = 85^\circ\ 10'$, and taking n successively as $0, \frac{1}{2}, 1, 2, \infty, -2, -1, -\frac{1}{2}$, the values given in Table III were calculated for $\phi_1, \phi_2, \theta_1, \theta_1', \theta_2, \theta_2'$. These values were then set out on the two stereographic diagrams shown in figs. 7 and 8.

Table III.—Spherical Polar Co-ordinates of Directions which are unextended during an increase of 30 per cent. in length of the Specimen.

n.	Unstrained material.			Strained material.		
	ϕ_1°.	θ_1°.	$\theta_1'^\circ$.	ϕ_2°.	θ_2°.	$\theta_2'^\circ$.
−2	−62·6	52·0	126·2	−63·3	37·4	141·4
−1	−44·2	57·8	125·9	−42·1	44·3	138·0
−½	−25·7	67·7	122·5	−22·7	56·5	130·5
0	0	87·3	112·9	0	80·7	114·5
½	25·8	111·6	93·0	20·5	111·6	87·0
1	44·7	121·0	81·1	35·2	124·7	71·4
2	63·4	127·5	69·4	51·6	135·3	56·1
∞	90·6	131·0	58·1	81·8	143·7	41·8

It was noticed that the points of these diagrams appeared to lie on two great circles, and on placing the tracing paper on which the diagrams were drawn on the stereographic net it was found that two great circles could be described to pass very close to all the points. These great circles are marked in figs. 7 and 8. The unextended cone is therefore a degenerate form, consisting of two planes in this case.

This is very different from the form which would arise when a bar was stretched which consisted of a number of small crystals orientated at random. The "unextended cone" would in that case be a circular cone whose axis coincided with the axis of the specimen.

Before proceeding further it may be well to notice the nature of the strain which causes the unextended cone to take the form of two planes. All figures described on either of these planes are unchanged both in dimensions and in shape by the strain; that is to say, the strain does not affect the relative positions of particles in either of them. Since the density of the material is practically unchanged by the strain, the distance apart of any two planes of particles which are parallel to either of these planes is also unchanged by the strain. The strain brought about by the extension of 30 per cent. can therefore be regarded as being due to a simple shear parallel to either of the two unstretched planes.

At a later stage we shall see that the X-ray analysis shows that the orientation of the molecules remains constant with respect to one of them, but not with respect to the other. From the external measurements taken at only two stages of the extension, however, we cannot distinguish between them and say that the shear is due to slipping on one plane rather than on the other.

Application of the Method to other Stages of the Extension.

The same method was applied to the other stages in the extension. It was found that in the cases 0–10, 0–15, 0–20, and 0–40 per cent. extension, the same result was obtained as in the case of 0–30 per cent., which has already been studied. In each case the cone of unstretched directions was found to consist of two planes. In the case of the final extension 78 per cent., however, it was found that this was not the case. We shall therefore study first all the stages of the extension up to 40 per cent.

On comparing the stereographic diagrams showing the positions in the unstrained material of the planes which are unstrained at the various stages of the extension up to 40 per cent., it was found that one of the planes in each case consisted always of the same particles.

Table IV.—Spherical Polar Co-ordinates of Directions in unstrained Material which remain unstretched at four stages of the test, namely, 0–10, 0–20, 0–30, 0–40 per cent. extension.

n.	ϕ_1°.	θ_1°.				θ_1'°.			
		0–10.	0–20.	0–30.	0–40.	0–10.	0–20.	0–30.	0·40.
0	0	86·5	87·5	87·3	87·0	114	113	113	111
1	45	122·3	120·5	121·0	120·2	74·3	78·2	81·3	82·2
∞	90·6	132·5	131·3	131·3	130·0	48·8	53·0	58·0	60·2
−1	−45	57·2	58·5	58·2	58·5	133·3	129·2	125·8	123·2

In Table IV are given the spherical polar co-ordinates of eight directions which remain unextended at each of the four stages, 10, 20, 30, and 40 per cent. extension. It will be seen that the figures under heading θ_1, which give the position before straining of one of the unextended planes, are practically identical for all four stages of the test. The figures under the heading θ_1', which gives the positions of the other planes, vary according to the particular stage concerned. It appears, therefore, that if the shear which produces the extension be considered as due to slipping on the plane given by the columns under heading θ_1 (Table IV), this plane of particles remains undeformed and unextended throughout the whole extension from 0 to 40 per cent. If, however, the other planes determined by the figures under the heading θ_1' be considered as planes of slipping, these planes will consist of different particles at different stages of the extension.

It is evident that the former is a simpler physical conception, but one cannot be sure that it is the correct one till one studies the directions of the crystal axes. Before proceeding to do this, however, there is still some more information to be got out of the measurements of the external shape of the specimen, the direction of the shear on the slipping plane can be determined.

Direction of Slipping on Unextended Planes.

Let us now consider how the unextended planes would be situated in the case of a simple shear on a given plane of particles. Let ABCD (fig. 9) be a rectangle in the plane perpendicular to the plane of slip, and let AD be in the plane on which the slipping takes place. Let the sheared position of BC be B′C′. All particles in the line BB′CC′ shift through distance equal to BB′. One of the unextended lines of particles is evidently AD. The other is found by taking two points E, E′ at distances equal to ½ BB′ on opposite sides of B. The line EA in the unstrained material moves to E′A in the strained material and AE = AE′, so that AE, AE′ are the traces of the second unextended planes.

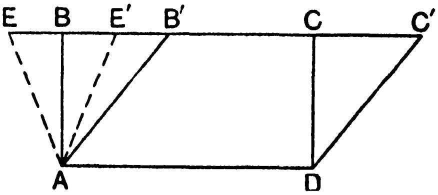


Fig. 9.—Section of material shearing on a slipping plane.

From these considerations it will be seen that:—

(*a*) The direction of shear is in the line at right angles to the line of intersection of the two unstretched planes.

(*b*) In a finite shear, due to slipping on one plane, the strained and unstrained positions of the other plane of particles, which is undeformed, lie at equal angles on opposite sides of the plane, which is perpendicular to the plane of slip and the direction of slip, *i.e.*, the plane whose trace is AB. If the slip be regarded as being measured by tan BAB′, fig. 9, this is equal to 2 tan BAE, *i.e.*, where $\frac{1}{2}\pi-$BAE is the angle between the two undeformed planes.

(*c*) The only plane which remains undeformed during the whole course of a finite shear is the plane of slip.

To determine the direction of shear, therefore, it is necessary to find the line of intersection of the two planes which constitute the "cone of unextended directions." In the present case this was accomplished by constructing a stereographic diagram in which the plane of slipping, *i.e.*, the plane which was undistorted throughout the extension from 0 to 40 per cent., was represented by the outer circle of the diagram, the normal to the plane of slipping being the centre.

In the case of the figure corresponding with the unstrained material this amounts, in the present case, to rotating the figure through 40° about the line of intersection of the slip plane with the plane perpendicular to the axis of the specimen. This operation can easily be performed with the stereographic net. In the case of the 30 per cent. extension diagram, for instance, the net is placed with its poles on the points L, M (fig. 7). The new positions of the points of the diagram are then found by moving them 40° round on a latitude circle of the net.

In the case of the figure for the strained material the same procedure is adopted, but in this case the angle through which the diagram has to be rotated is different for different extensions. In each case the boundary circle of the figure represents the same plane of particles, and, in order to make it easy to superpose the different figures, the directions of the lines of intersection of the slip plane with the faces and with the axial planes which are parallel to the diagonals of the cross-sections are marked as dots on the edge of the stereographic diagram for each extension. The fact that these dots can be superposed with diagrams obtained from measurements taken at different stages of the extension is confirmatory evidence that the slip plane is undistorted.

In fig. 10 is shown the position in the unstrained material of the second set of planes, which are undistorted at four stages of the extension, namely, 10, 20, 30 and 40 per cent. In fig. 11 is shown the positions of three of these same planes in the strained material.* It will be seen that in both cases all the great circles which represent the second set of planes cut the boundary of the diagram near the same two points. The direction of slip, which, as can be seen from consideration (*a*) above, is represented by the point on the boundary at 90° from these points, is therefore *practically the same for all the stages of the test up to* 40 *per cent. elongation of the specimen.*

It will be noticed that the relation between figs. 10 and 11 is what we should expect from consideration (*b*) above. The fact that the dotted lines in fig. 10 correspond with the full lines in fig. 11 and *vice versâ*, shows that the second set of planes make equal angles on opposite sides of the plane perpendicular to the direction of shear.

---

* The position of the second undistorted plane for 10 per cent. extension has not been shown in this figure because it confused the central part of the diagram.

To summarise, we have shown that, up to 40 per cent. elongation, the distortion of the specimen is due to simple shear in one direction on a certain plane. We have determined the plane and the direction. It now remains to determine the relation between this plane and the crystal axes.

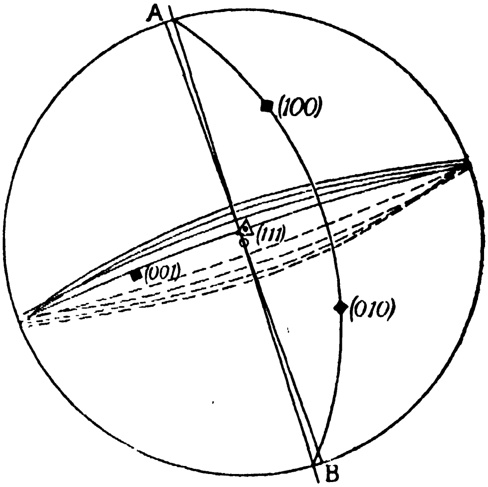


Fig. 10.—Projection on to the plane of slip of crystal axes and of second unstretched planes for extensions 0–10, 0–20, 0–30, 0–40 per cent. *Unstretched* material.

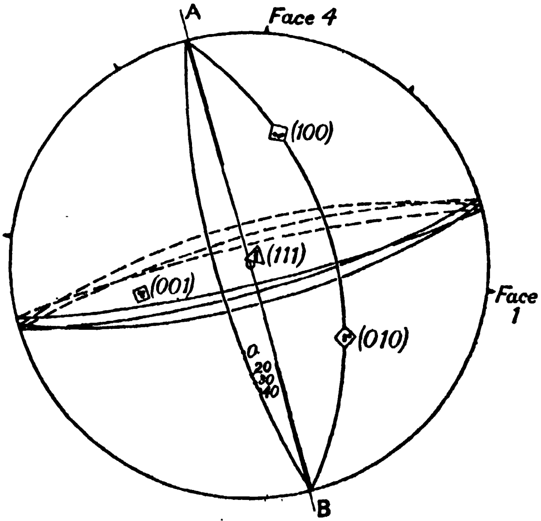


Fig. 11.—Positions in *strained* material of the axes and planes shown in fig. 10. The positions of the axis at 0, 20, 30 and 40 per cent. extension are also shown.

Relation between Plane of Slip and Crystal Axes.

The co-ordinates of the crystal axes were measured by Dr. Müller and Miss Elam at each stage during the extension. They are given in Table V.

Table V.—Spherical Polar Co-ordinates of (111) and (100) Axes at various Stages of the Test.

e.	θ_{111}°.	ϕ_{111}°.	θ_{100}°.	ϕ_{100}°.	i°.
1·000	131	269·3	79·0	262·2	52·3
1·053	127·4	270·2	75·0	263·8	52·5
1·110	126·0	270·3	73·3	263·2	53·2
1·161	123·5	270·9	70·2	263·8	53·6
1·200	123·2	270·5	69·7	263·2	54·0
1·304	118·9	221	69·5	262·2	50·2
1·404	118·2	272·2	66·0	260·5	53·4
1·623	113·9	274·0	62·0	260·6	52·8
1·785	114·7	272·8	62·4	257·7	54·5

NOTE.—i is angle between (111) and (100) axes. It should be 54·7°

It will be seen that one of the (111) planes and one of the (100) planes were determined. In order to test the accuracy of the determinations, the angles between these two directions, which should be 54° 44′, were worked out. These angles, which are represented by the symbol i in Table V, are given in the last column. It was found that in some cases there was an error of 2°, and in one of them an error of 4°.

When the positions of two of the axes are found, it is possible to use our knowledge of the type of symmetry which the crystal possesses to determine the positions of the others, but when the measured angle between these two axes is not what we should expect from the known symmetry of the crystal, there is a considerable degree of uncertainty in placing the other axes. In placing the axes, I have accordingly adopted the following system :—I have taken the great circle through the two measured axes as being correct. I have then taken the mid point between the two measured axes as correct, and have shifted both axes by equal amounts in opposite directions along the great circle joining them till their distance apart is 54° 44′. This is equivalent to shifting each of them through an angle equal to half the error in the determination of the angle between them. It is then possible to find the positions of the remaining axes. Such a proceeding is very arbitrary, but something of the kind is necessary, and any system adopted must be arbitrary. The maximum angle through which either of the axes is shifted by this scheme of averaging is 2·2° in the case of the 30 per cent. extension. The shifts in all other cases are less than this, the average being 1·0°.

Fortunately, the distortion of the material is so great that very large changes in the orientation of the crystal axes relative to *some* planes of particles must take place.

To find the relative positions of the axes and the slip plane, the particular triad (111) axis which was measured and the three tetrad (100) axes have been marked on the stereographic diagrams (figs. 10 and 11). Fig. 10, which applies to the unstrained material, of course, has only one set of axes marked on it, but fig. 11 has a set of axes for each stage of the extension.

On looking at these diagrams the following facts will be noticed :—

(1) The (111) axis is close to the pole of the diagram. That is, the plane of slip is nearly coincident with a (111) plane of the crystal.

(2) The positions of all the axes are fixed relatively to the plane of slip during the whole course of the distortion, though, of course, the lines of particles which originally occupied positions along the axes move through very considerable angles with respect to it. The variation in the position of the normal to the (111) plane, for instance, is only about 4°.

(3) The great circle through two of the (100) planes is very nearly parallel to the direction of slip. The angle which the line of intersection of the (100) plane containing these two (100) axes and the plane of slip makes with the direction of slip is less than the margin of experimental error.

If the planes represented by the symbols $(\pm 1 \pm 1 \pm 1)$ be regarded as forming a regular octahedron, the slip is then on a plane which is very nearly parallel to one face, and is almost exactly in the direction of one edge.

The fact that the position of the slip plane determined by external measurements is so close to the (111) plane determined by X-ray analysis makes one suspect that they really coincide, and that the difference between their measured positions is due to experimental error. This idea is confirmed by the fact that the direction of slip is so exactly parallel to one of the three principal lines of atoms in the (111) plane. It may be, however, that the difference between the measured positions of the slip plane and the (111) plane is due to some slight slipping on some other plane. During the test, two sets of slip bands appeared on the faces of the specimen. It is hoped to settle this point in another test, in which improved methods of measurement will be used.

In the remaining part of this paper we shall assume that the slip was actually on the (111) plane, and shall discuss some possible consequences of this assumption.

Changes in Orientation of Specimen relative to Crystal Axes during Extension.

The effect of a shear in any direction on a fixed plane of particles is to move any given line of particles in such a way that it remains parallel to the plane containing its original direction and the direction of slip. In fig. 11, for instance, the positions of the axis of the specimen relative to the slip plane at successive stages of the extension is shown. These are marked 0, 20, 30, 40. It will be seen that a great circle can be drawn to pass through them and through the direction of slip.

As the point representing the direction of the axis moves along this circle the angles between this direction and all the four octahedral planes vary continuously. After a time, it seems likely that the axis of the specimen will move into a position in which it makes an angle with one of the other octahedral planes which is equal to the angle it makes with the particular octahedral plane on which the slipping is taking place. At this stage the force tending to produce slipping on one plane is equal to that tending to produce slipping on the other. It is possible, therefore, that at this stage slipping may take place on both planes. Possibly this is the reason why the cone of unstretched directions in the case of the 78-per-cent. extension no longer consists of two planes.

To examine this point, first turn the axes round so that one of the cubic or tetrad axes is in the centre of the figure. The two remaining cubic axes will then be found on the circumference, and the four octahedral axes will be disposed symmetrically round the central cubic axis.

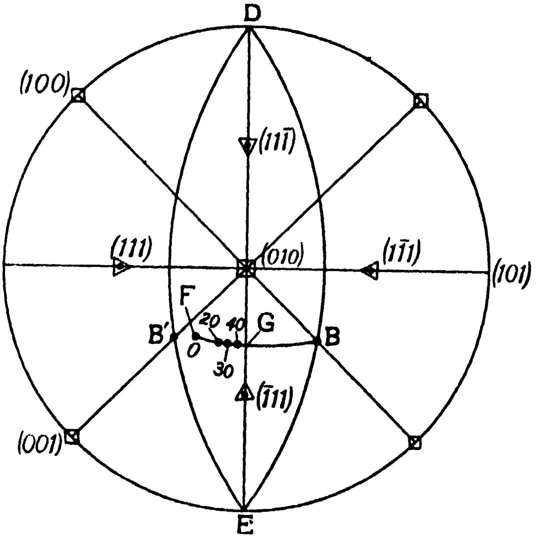


FIG. 12.—Figure showing position of axis of specimen relative to crystal axes at 0, 20, 30, 40 per cent. extension.

In fig. 12 the axes have been rotated so that the cubic axis denoted by (010) in fig. 11 is in the centre. The slip plane is then represented by the great circle DBE (fig. 12), and the point B represents the direction of slip, corresponding with the point B in fig. 11.

We have already shown that the points representing the axis of the specimen lie on a great circle through B. An arc of this circle is shown in fig. 12 as FGB, where G is on the line DE which represents the (101) plane.

As the slipping on the plane DBE proceeds, the point representing the axis of the specimen moves along the arc FGB till, at about 50 per cent. extension, it gets to the point G. At this point, G, it is in such a position that it is equally inclined to the two octahedral planes DBE, DB′E, which are perpendicular to the two axes (111) and $(1\bar{1}1)$.*

It should now be noticed that up to this stage of the extension the slipping has always been along that one of the twenty-four crystallographically similar possible directions of shear which has the maximum component of shearing force tending to set it going. If this condition continues to hold, then directly the point representing the axis has moved beyond G along its path FGB, the force tending to produce a slip on the plane DB′E in the direction of B′ is greater than the force tending to continue the slip on DBE. Slipping will then begin on DB′E, and the axis will start moving along a great circle towards B′.

It can easily be seen that the result of successive slips towards B and B′ will be that the point representing the axis will move along the line DGE towards the point where the great circle through B and B′ cuts it. At this point no further movement of the axis will take place, however much the specimen extends.

These considerations lead to the prediction that there will be a tendency for the axes to orientate themselves, so that two of the octahedral planes make equal angles of 61° 52′ with the axis of the specimen, while their line of intersection makes an angle of 54° 44′ with the axis.

It is hoped to examine this point later, when the distortions and crystal axes of more specimens have been determined.

Verification that Slipping in the Last Stage of the Distortion occurs simultaneously on Two Octahedral Planes.

Though no very satisfactory method has yet been devised for determining the distortion due to simultaneous slipping on two planes, it is possible to apply the methods used in the early part of this paper to test whether the distortion of the specimen at its breaking point (78 per cent. extension) could have been produced by slipping on the two octahedral planes which are at equal angles to the axis when the extension is about 50 per cent.

---

* No distinction is made here between the planes $(1\bar{1}1)$ and $(\bar{1}1\bar{1})$.

The cones of unextended directions in the distortion of the specimen during its extension from 40 to 78 per cent. were first determined by the equations (1), (2), and (3). The two positions of this cone are plotted on the stereographic figures 13 and 14. Fig. 13 shows that the position of the cone in the material at 40 per cent. extension, while fig. 14 shows its position at 78 per cent. extension. In both these figures the centre spot represents the

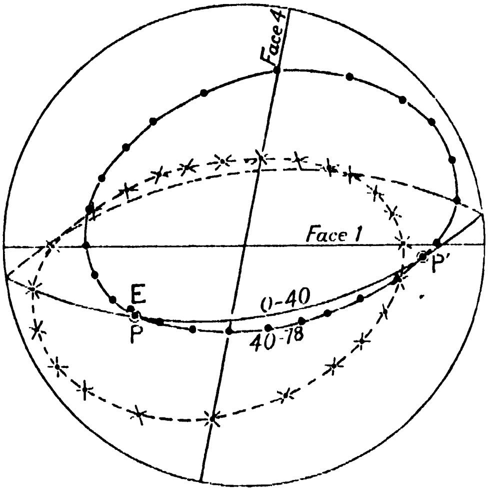


Fig. 13.—Positions in material extended 40 per cent. of unstretched cone for extension from 40 to 78 per cent., and of slip plane for extension 0 to 40 per cent.

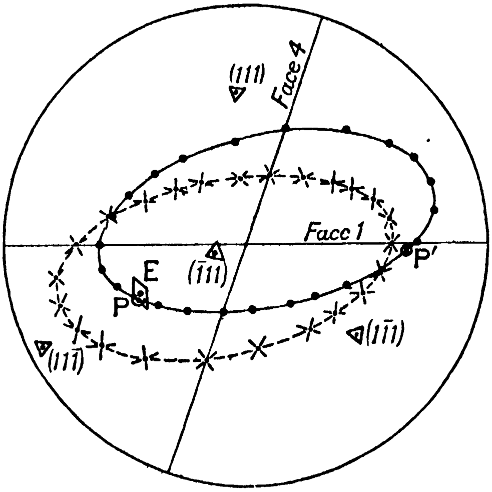


Fig. 14.—Position in material extended 78 per cent. of unstretched cone shown in fig. 13.

axis of the specimen, as it does in figs. 7 and 8. It will be seen that the cone is now very different from the pair of planes which were characteristic of the first 40 per cent. of extension. The distortion is, therefore, no longer due to slipping on one plane.

We have seen that during the last stage of the extension it seems likely that slipping might take place on the second octahedral plane, DB′E (fig. 12). If this is actually the case, then there is one line of particles in the specimen which will not stretch during the whole extension from 0 to 78 per cent., namely, the line of intersection of the two octahedral planes (111) and $(1\bar{1}1)$. The position of this crystal axis, which is represented by the symbol $(10\bar{1})$, can be found from the crystal measurements. It is a line which makes 45° with the plane (100) and 90° with the plane (111). Using the spherical polar co-ordinates (θ, ϕ), previously used in Table V for giving the positions of the crystal axes, it was found that the co-ordinates of this crystal axis were $(\theta = 56{\cdot}5^\circ, \phi = 209{\cdot}6^\circ)$ in the material at 40 per cent. extension, and $(\theta = 51{\cdot}5^\circ, \phi = 204{\cdot}8^\circ)$ in the material at 78 per cent. extension.

To find the directions which were unextended during the extension 0–40 per cent., and remained unextended at 78 per cent. extension, the positions in the material at 40 per cent. extension of both the slip plane for the extension 0 to 40 per cent., and the " unextended cone " for the extension 40 to 78 per cent., were marked on the same diagram (see fig. 13). The points P and P′ where these cut then represent the directions required. It is clear that one of them must have remained unextended during the whole of the stretching from 40 to 78 per cent., whereas the other is merely a line of particles which happens to be unextended at the particular stage, 78 per cent. of extension; but it is not possible to prove that this is the case, because there are no reliable measurements available for stages intermediate between 40 and 78 per cent.

One of these two directions should therefore, if our theory is correct, coincide with the diad axis whose position in the 40 per cent. material is $(\theta = 56{\cdot}5^\circ, \phi = 209{\cdot}6^\circ)$.

On measuring up fig. 13 it was found that the co-ordinates of P were $(\theta = 57^\circ, \phi = 213^\circ)$, and the co-ordinates of P′ were $(\theta = 73^\circ, \phi = 355{\cdot}5^\circ)$.

The diad axis $(10\bar{1})$ is marked at the point E (fig. 13). It is clear that it is very near to its predicted position.

Verification of Relation between Unstretched Lines of Particles and Crystal Axes in Material at Breaking Point.

The positions in the 78 per cent. material of the unextended lines of particles P and P′ are marked on the diagram (fig. 14). The co-ordinates of the diad axis, which lies on the intersection of the (111) and the $(1\bar{1}1)$ planes, were calculated from the crystal measurements given in the last line of Table V; they are ($\theta = 51·5°$, $\phi = 204·8°$). This diad axis is marked on the diagram (fig. 4) at the point E.

The co-ordinates of the point P (fig. 14) were determined by measuring the diagram; they are ($\theta = 53°$, $\phi = 206·6°$).

The co-ordinates of the diad axis and the unstretched lines P are given in Table VI.

Table VI.—Spherical Polar Co-ordinates of Diad Axis and unextended Direction.

Extension of material	40 per cent.		78 per cent.	
Co-ordinates	θ	ϕ	θ	ϕ
Diad axis	56·5°	209·6	51·5°	204·8°
Unextended direction	57°	218·0	53°	206·6°

The agreement between these two is striking, when it is remembered that the diad axis was determined in each case by X-ray analysis of the crystal at one stage only, while the direction of the unextended line of particles was found, using only external measurements of the specimen at a series of successive stages of the distortion.

It will be noticed that this agreement depends only on the fact that the slipping is on the two octahedral planes (111) and $(1\bar{1}1)$. If the slipping on the two planes were of equal amounts (as contemplated on p. 665 above) we should expect the two octahedral planes on which the slipping takes place to remain at equal angles to the axis of the specimen during the last part of the stretching. On calculating these angles from the X-ray measurements it was found that they were 66·1° and 63·5° in the material at 60 per cent. extension, and 65° and 61·5° in the material at 78 per cent. extension. The limiting value should, according to the theory, be 61° 52′. It will be seen that the agreement is as good as could be expected in view of the uncertainty of the X-ray measurements. The symmetry of the octahedral planes with respect to the axis of the specimens during the last stages of the extension is shown in fig. 14, where the four triad axes, which are normal to the four octahedral planes, have been marked as spots surrounded by triangles.

It will be seen that the two planes (111) and $(1\bar{1}1)$ are at nearly equal angles to the axis, and that the other two triad axes lie on a plane which very nearly passes through the axis of the specimen.

The cone of unextended directions for the extension 40 to 78 per cent. is not symmetrical with respect to this plane, because in the early part of this range of extension the slipping was only on the original slip plane, but if the external measurements at 62 per cent. extension had been satisfactory I should have expected the cone of unextended directions for the extension 62 to 78 per cent. to be symmetrical with respect to the (101) plane. In this way it is hoped that it will be possible to test whether there is an equal amount of slip on the two slip planes during the final stage of the extension. The fact that the (111) plane and the $(1\bar{1}1)$ plane are nearly equally inclined to the axis both in the case of 60 per cent. extension and in the case of 78 per cent. extension suggests that this must be the case, but it cannot be proved without further external measurements.

[*Note added, February* 8, 1923.—Since this paper was communicated, an account of some similar work by Mark, Polanyi and Schmid has appeared in the 'Zeitschrift für Physik,' December, 1922. The distortion of a zinc crystal is discussed. The method used is based chiefly on measurements of slip bands (which in the case of aluminium were found to be nearly useless as a basis for distortion measurements). The results obtained are similar to those described in the present work, but, owing to the fact that zinc has hexagonal instead of cubic symmetry, the complication introduced by the existence of a large number of crystallographically similar slip planes does not arise.]

Studies of Iridescent Colour, and the Structure producing it.—I. The Colours of Potassium Chlorate Crystals.

By LORD RAYLEIGH, F.R.S.

(Received January 1, 1923.)

[PLATES 6 AND 7.]

It is known that in the crystallisation of potassium chlorate from solution, crystals are occasionally found which show a brilliant iridescence. The subject was investigated by Stokes* and by my father, the late Lord Rayleigh.† The latter was the first to conclude that the character of the reflexion required that it should come from a large number of reflecting surfaces at regular intervals. These hypothetical surfaces he regarded as twin planes. Only by the postulate of a large number of surfaces can we account for the highly selective character of the reflexion, limited in many cases to a narrow band in the spectrum. (See Plate 6, fig. 1. The sodium and lithium lines are placed on the plate for comparison.)

He wrote (*loc. cit.*): "It would undoubtedly be far more satisfactory to speak of the periodic structure as a matter of direct observation, and it is to be desired that some practised microscopist should turn his attention to the subject. *Ex hypothesi* we could not expect to see the ruled pattern upon a section cut perpendicularly to the twin planes, as it would lie upon or beyond the microscopic limit. I have tried to detect it upon a surface inclined to the planes at a very small angle, but hitherto without success."

Recently examining my father's collection of these coloured crystals, I remembered his discussing with me, about 1912, the feasibility of seeing the structure, and I determined to make the attempt. The subject was one to which he often returned in his later years, and he was also very desirous of finding what were the peculiar conditions determining the formation of coloured crystals.

If we regard the stratified structure which is to be examined in section, as modelled, *e.g.* on an ordinary pile of glass plates, then the period for first order reflexion is $\frac{1}{2}\lambda$ which, as remarked in the quotation above, is on the microscopic limit. But by the use of polarised light, which I believe my father did not contemplate, it is possible advantageously to modify this condition: for by suitably orientating the section relative to the crossed nicols we can make the alternate twin strata dark. In this case the appropriate model is a pile of black and white glasses alternately, and the complete linear period is twice as great as before. This makes the whole difference for microscopic resolution.

---

* 'Roy. Soc. Proc.,' vol. 38, p. 174 (1885).

† 'Phil. Mag.,' vol. 26, p. 256 (1888).

Again, there is the possibility of structures equally regular but much more open than those giving coloured first order reflexions. Such structures would show reflexions of the second or higher orders.

Some experimenting was necessary to learn how to prepare for microscopic examination thin sections perpendicular to the plane of the tabular crystals. The material is very friable and will not bear rough handling. The crystal was cemented down with Canada balsam to a solid support, and a piece cut off with a fine saw. The balsam was removed with benzol, and the piece of crystal was then placed on edge on a glass slip, supported on either side by cleavage blocks of rock salt. The chlorate and rock salt were cemented to one another and to the glass slip with balsam of such consistency as to harden on cooling. The top surface was ground away on fine emery paper until flush, and the surface was then polished by licking it with the tongue. When this had been done, the whole was inverted, cementing the polished surface to the glass slip. It was then ground down from the other side as far as was considered safe, and polished by again licking it. It was further cautiously licked to diminish the thickness by solution without the violence of abrasion; when bright interference tints appeared locally under polarised light it was a sign that the danger point was being approached, and holes were apt to develop. A cover-glass was then cemented on, using balsam in benzol, and the section was ready for examination. I have no definite measure of the ultimate thickness, but it is probably greater than the 0·025 mm. usually aimed at in making sections of hard rocks. Much thinner sections would be desirable if the mechanical difficulties of making them could be overcome.

Rock salt was used for the supporting blocks, because its hardness is comparable with that of the chlorate, and thus the two can be conveniently ground away together. Its use is quite comparable with the use of paraffin wax for embedding biological material which is to be sectioned.

In selecting crystals for examination, I usually avoided those which showed a strong monochromatic reflexion, because it appeared probable from the theory that these would be difficult to resolve microscopically. Other crystals are found showing pale tints and giving a complex spectrum, and these were more likely to show structure which might admit of, at all events, partial resolution. This anticipation was confirmed. Plate 6, fig. II, is a photograph of a section of such a crystal, using polarised light to develop a contrast between the successive twin planes. The cross lines indicate the planes of the two nicols. The magnification is 360 diameters. A $\frac{1}{4}$-inch dry objective was used.

I do not put forward this photograph as a complete picture of the structure of the crystal such as would allow of a complete prediction of the character of the reflexion.

As Mallock and others have emphasised, in examining a fine periodic structure, it is most desirable that the thickness of the section should not exceed the linear period of the structure. The sections here used by no means satisfy this condition, and I doubt the feasibility of satisfying it, at any rate in a section of any considerable area, where such brittle material is concerned. Where, as in the present case, the section is thicker, the parts which lie out of the focal plane may cause confusion. Subject, however, to this limitation, the photograph shows clearly enough the existence of the multiple reflecting planes, and thus satisfies the purpose originally in view. The details of what is seen in such a section depend very much on the exact conditions of lighting. The clearness with which the stratifications come into view depends on circumstances which I have not altogether succeeded in controlling. It often makes a difference if the polariser and analyser (which in the instrument used are geared to rotate together) are rotated through 90 degrees, thus interchanging their planes of vibration. The known polarising action of gratings perhaps intervenes.

The next photograph (Plate 6, fig. III) shows the best resolution I have obtained on any of these fine structures. The magnification is 3200 diameters, thus the complete linear period is 0·81 μ,[*] and if we assume the alternate strata to be equal in thickness, the half period, representing the distance from one reflecting plane to the next, is 0·405 μ. The equivalent optical length in air would be about 0·61 μ.[†] This is one wave-length of red light, and, therefore, corresponds to a second order reflexion of this kind of light near normal incidence. There are other stripes visible on the negative which will probably not come out in the reproduction. Why they are so much fainter I do not know.

The spectrum of the reflexion from the crystal whose structure is shown in Plate 6, fig. II, is very complex. A photograph of it is reproduced on Plate 6, fig. IV, taken by means of a pointolite lamp, and a 3-prism spectrograph with camera lens 3 inches focus. The sodium and lithium lines have been put on for comparison, and the former is well resolved as a doublet on the negative, and may perhaps be seen as such on the reproduction, though the line is faint. I have also marked the approximate position of the green helium line 5016.

---

[*] $\mu = \frac{1}{1000}$ mm.

[†] The refractive index of potassium chlorate is taken as 1·5. I have not found any measurement in the books which are to hand.

Some of the lines from the chlorate crystal will be seen to be so narrow that their breadth is not more than half the interval between the D lines.

High resolving power, or discrimination between one region of the spectrum and another very near to it, requires for the chlorate reflexion conditions analogous to those applicable to diffraction spectra. We must either have a large number of planes, or alternatively, a reflexion of high order. If we supposed every one of the lines in this complex spectrum to result from a separate assemblage of planes, each assemblage giving a first-order reflexion, with one-wave relative retardation, then a number of planes would be postulated which it would be impossible to find room for in the thickness of the crystal. It is not worth while to enter upon elaborate computations. In order to resolve the D lines in the first order we require 1000 reflexions, so that at normal incidence it is necessary to have 500 wave-lengths *in the chlorate*, or, say, a thickness of about 0·2 mm.

There would only be room for 6 such assemblages, corresponding to 6 lines in the spectrum, in the entire thickness of the crystal, which is a little over 1 mm. even if it were entirely filled with twin planes, whereas the spectrum contains perhaps 100 lines of varying fineness but comparable with this, even in the limited region between the red and green.

We are compelled, therefore, to fall back on the alternative of high relative retardation. It is true that this requires as great thickness for a given resolving power, as if a large number of planes conspire to give a first order reflexion. Thus, if there are two planes only in question, to produce as fine a line as those we have been discussing would require the 500th order of interference, and thus the thickness would still be 0·2 mm. But there is this difference, that a great number of lines of this fineness would be accounted for, representing interferences of orders a little higher or lower; moreover, the available space in the crystal can be used many times over, as it may form part of the interval between other pairs of planes than the pair first considered.

The particular portion of the section here shown (Plate 6, fig. II) is only about $\frac{1}{14}$ part of the whole thickness. The section generally may be described as showing a single crystal of one orientation, crossed by several bands of complex twinning. The picture (Plate 6, fig. II, × 360 diameters) shows the most important of these individual complex bands, with the untwinned crystal (dark) on either side of it. Its breadth, about 0·08 mm., is not enough to give the resolving power exhibited in some of the finest spectrum lines, and we are thus compelled to admit that interference between the planes of this group and planes belonging to other groups must be concerned.

A complete analysis of the spectrum has not been made, and would perhaps hardly reward the trouble. The appearance of regular groups of lines confined to a limited range of the spectrum is to be accounted for as follows. A group of a moderate number of reflecting planes, say 20, regularly spaced like the group in Plate 6, fig. III, will limit the spectral range transmitted to, say, 500 angstroms, or in the case of a second-order reflexion, to 250 angstroms. Now if two such groups of about the same period occur in the crystal, separated from one another by a mean distance of 0·5 mm., we shall get a bright region of 500 angstroms cut up into lines of about 12 angstroms broad, by the high interference of the two groups each regarded as a separate whole.

This may be taken as a simple example of the kind of way in which the complex reflexion spectrum is to be explained, without making excessive demands on the amount of space available in the thickness of the crystal. The possible complexities arising from interferences of different orders between different groups of periodically spaced planes are obviously very great. There seems to be no reason for surprise at the character of the reflexion. Periodically repeated groups of lines will be recognised, and also narrow lines and broad bands, not obviously related to other parts of the spectrum.

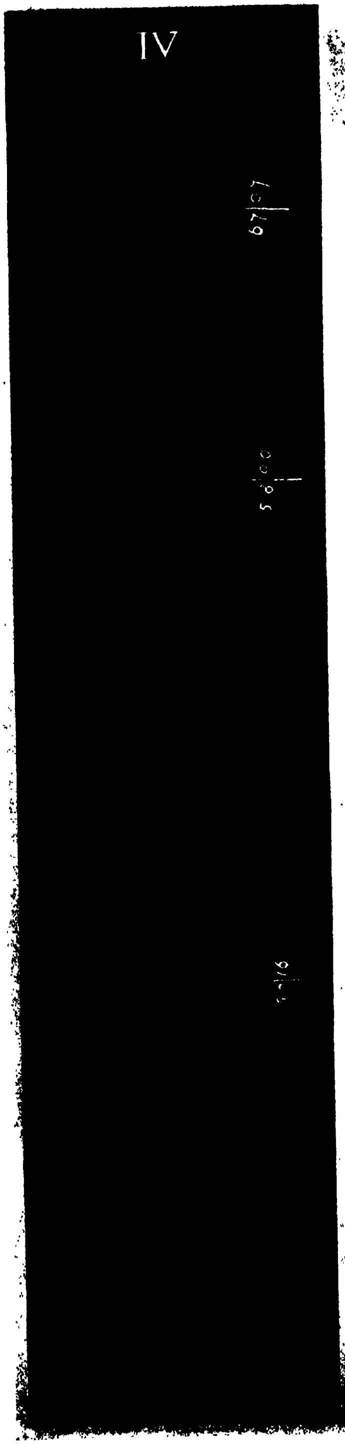
Madan* showed that ordinary chlorate crystals, not coloured, when heated to a temperature of about 250° C., gave, on cooling, a silvery white reflexion, which, as my father showed, has many of the same properties as the coloured reflexions already mentioned. He wrote: "The only difference I should suppose to exist between the constitution of these crystals and the coloured ones is that in the former case the alternations are irregular, and also probably more numerous."

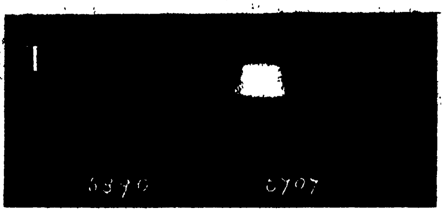
I thought it would be of interest to examine these crystals by the methods already described, for comparison with the others. Plate 7, fig. V, shows in section a thin ordinary flake-like crystal of chlorate as obtained by a laboratory crystallisation. It will be seen that the material is homogeneous, except for two thin isolated twin layers. A portion of the identical crystal was treated by Madan's method, heating it on an iron plate short of the melting point, and allowing it to cool. A section made from the portion thus treated is shown in Plate 7, fig. VI (× 360 diameters). The change of structure is (to me at least) very wonderful.

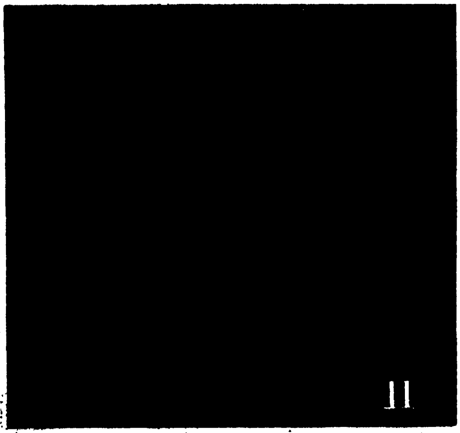
The general aspect of this section certainly shows much less regularity than the crystals obtained from aqueous solution. The latter have a perfectly straight and parallel structure, but here there is, in places, a curious crossing

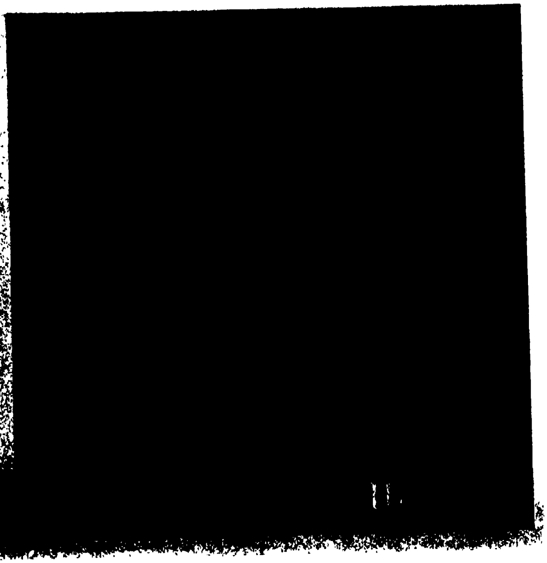
---

* 'Nature,' vol. 34, p. 66 (1886).


IV
6707
5890
5176


I
5890
6707


II


II

and interlacing of fibres, which must represent a very complex structure in three dimensions.

In other places, however, assemblages of several parallel equidistant planes can be made out, at least on the original negative. The spectrum of the reflexion (Plate 7, fig. VII) shows much the same kind of irregularities as the section of the crystal itself does under the microscope. Want of parallelism in the reflecting planes has its counterpart in the lack of straightness of the spectrum lines. These lines are not so fine as some of those seen in Plate 6, fig. IV, but they are in some cases not more than 10 angstroms in breadth, and no doubt this is due, as before, to high interference from two separated groups of parallel planes. The appearance of the section hardly allows us to postulate single groups of equidistant planes numerously enough to give the observed resolving power, and there is the difficulty of finding room in the thickness of the crystal, in this case as in the previous one.

Summary.

The structure of the iridescent potassium chlorate crystals investigated by Stokes and the late Lord Rayleigh is examined microscopically.

The periodic twinned structure inferred by the latter is clearly shown in the photographs taken under the microscope with polarised light.

Some crystals have exceedingly complex structure, showing many groups of evenly spaced twin planes, and a very complex reflexion spectrum. It is shown that this results from high interference from terminal layers situated a considerable distance apart.

Chlorate crystals, giving a silvery reflexion, were obtained by Madan, who heated the ordinary colourless crystals to about 250° C. The complex twinned structure induced by this treatment is investigated, and photographs of the structure of the crystal and of the reflexion spectrum are reproduced. There are corresponding irregularities in each resulting from want of flatness in the twin planes.

DESCRIPTION OF PLATES.

PLATE 6.

Fig. I. Spectrum of the reflexion at nearly normal incidence from a strongly coloured crystal (simple type). Sodium and lithium lines for comparison.

Fig. II. Group of terminal planes seen in section under polarised light, × 360 diameters. The crystal contains other groups parallel to this one, and separated from it by uniform crystalline material.

Fig. III. Group of twin planes showing good resolution, × 3200 diameters. Note that 12 of the planes are nearly, though not quite, equidistant.

Fig. IV. Spectrum of the complex twinned crystal shown (partly) in fig. II.

PLATE 7.

Fig. V. **Section of an ordinary white chlorate crystal chosen at random, under polarised light, ×360 diameters. Note the two isolated twin strata.**

Fig. VI. **Section of a portion of the above (No. V) after heating to above 250° C. Note the complex twin structure produced. ×360 diameters.**

Fig. VII. **Spectrum of the reflexion from No. VI.**

Studies of Iridescent Colour, and the Structure producing it.— II. Mother-of-Pearl.

By LORD RAYLEIGH, F.R.S.

(Received January 1, 1923.)

[PLATE 8.]

§ 1. *Introduction.*

The colours of mother-of-pearl were studied by Brewster* and he recognised that the colours were to be classed under two heads—what he called transferable and non-transferable colours. The former could be transferred to gelatine or other soft substances, employed to take a cast of the mother-of-pearl surface, and in this way it was clearly shown that the colour was due to diffraction by grooves on the surface.

The non-transferable colours on the other hand were classified with the colours of thin plates, due to the stratified structure of the shell. The grooved surface was naturally explicable as resulting from the out-crop of these strata.

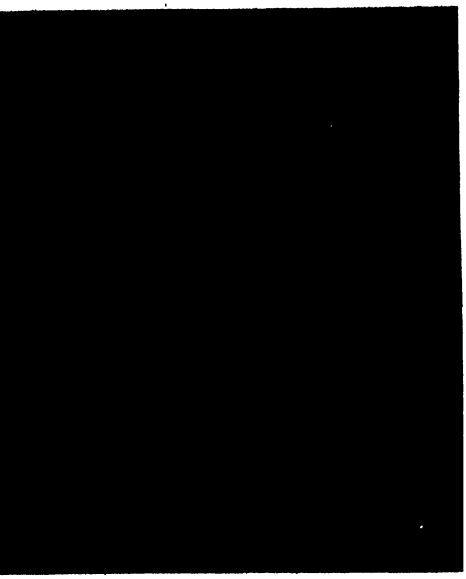
Brewster apparently did not recognise the peculiar monochromatic character of the reflexion from a series of equally spaced parallel plates, distinguishing it from the ordinary colours due to a single thin stratum. This was first enunciated by my father, the late Lord Rayleigh, in connection with the iridescent chlorate crystals (see the preceding paper), but it can be shown that the same action occurs in mother-of-pearl. This observation was first made by A. H. Pfund.† His method was to study by means of a thermo-couple the reflexion in the red and infra-red parts of the spectrum. Before becoming acquainted with his work I made a short study of the subject by means of the microscope and the spectrograph. The results are con-

---

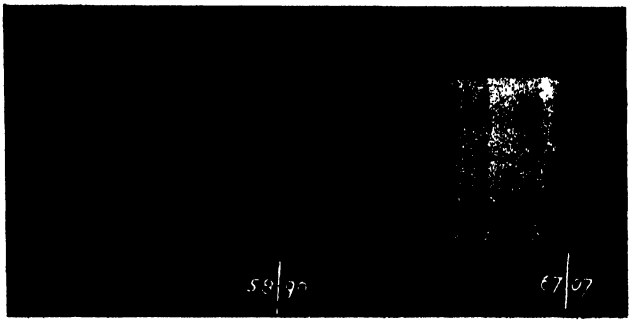
* For a summary see his 'Treatise on Optics,' p. 137 (1853).

† 'Journ. Franklin Inst.,' vol. 193, p. 453 (1917).

V



VI



VII

firmatory of his, though reached by a different method, and they seem worthy of brief record.

§ 2. *Transferable Colours.*

In Plate 8, fig. I, is shown a photograph by transmitted light of a selected area from the outer (convex) part of the nacreous layer of the shell of a pearl oyster. The magnification is 80 diameters. In preparing this section the rough external coating was broken off, and the nacreous layer ground away from the concave side till the section was thin and transparent. The external convex surface of the nacreous layer was not touched, except to apply the necessary cement, which was removed afterwards. The lines on the photograph, which correspond very closely with Brewster's drawing, show the edges of the successive layers, which overlap after the manner of a pack of playing cards which have been given a shearing displacement. These lines are at the rate of about 880 per cm., or 2200 to the inch.

This specimen shows curious peculiarities as regards the relative intensity of the diffraction spectra. If the observer holds it before his eye in a certain position, with the light behind him, he sees by reflexion 4 orders of spectra, the intensity diminishing with increasing order. These are on the right of the rather faint central image. The first order only can be seen on the left. If the observer turns to view the light by transmission, without moving the grating relative to the eye, the first order on the right is now very much brighter than the central image, or than any of the other spectra; of these latter, the third order on the right ranks next, and the second order on the right, together with the first on the left, rank after it.

Such a natural grating, formed of mother-of-pearl, may be regarded as a model of Michelson's echelon. There are a series of steps of graduated retardation, and the light is nearly all thrown into one order. Here it is the first order. In the artificial echelon it is a very high order. Unfortunately the steps of the natural echelon are not uniform enough to make it of any use for optical work.

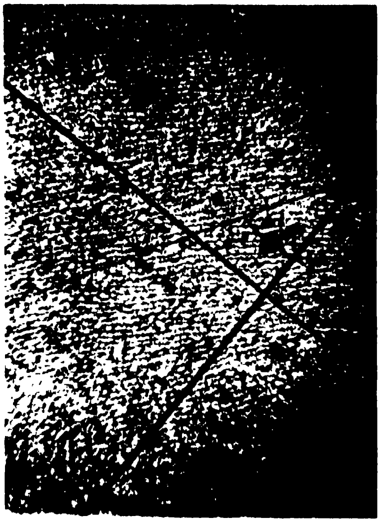
§ 3. *Non-transferable Colours.*

My observations on these colours were chiefly made on the brightly coloured nacreous layer of what is popularly called an *ear shell* The specimen gave non-transferable colours only. A section was cut transversely to the layer, and examined under polarised light. When the light was carefully adjusted, the stratified structure giving rise to the coloured reflexion came out with satisfactory distinctness. (Plate 8, fig. II, × 400 diameters.) The material is tough, and leaves a considerable residue of horny matter when treated with hydrochloric acid. It appears probable that the strata consist of calcite and horny matter alternately.

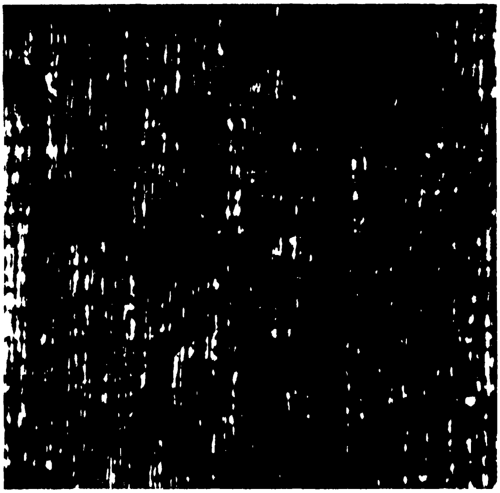
By examining the reflexion from the inside of this shell with a pocket spectroscope it is readily made out that the colour from any particular area is very much localised in the spectrum. At the same time there is considerable irregularity and change of colour in different parts of the surface. This is explicable in the main by local variations in the linear period. General want of flatness is of minor importance.

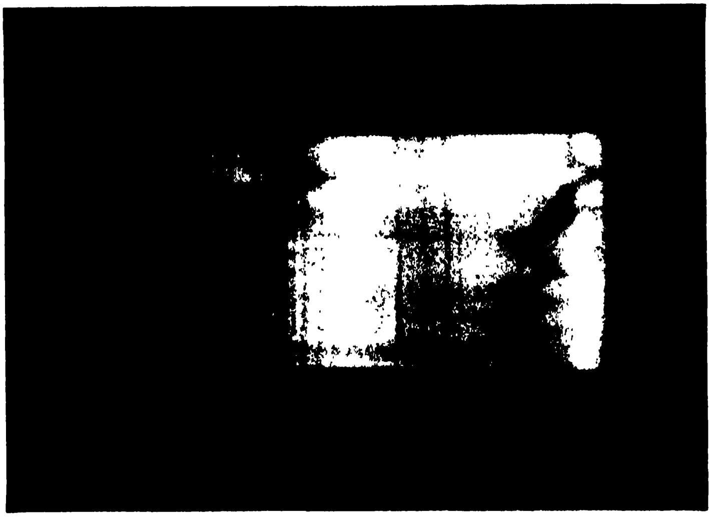
The spectrum of iridescent chlorate crystals is advantageously examined by reflexion, partly because in any case the characteristic phenomena do not appear at strictly normal incidence: thus the specimen must be sloped. This is a consequence of the nature of the reflecting planes in the chlorate. Mother-of-pearl, on the other hand, does not require sloping, and if we examine it by transmission the specimen can be put right up to the slit, and its parts differentiated. But there is another reason as well in favour of this course. Mother-of-pearl has a general cloudiness which scatters light of all wave-lengths, and thus dilutes the purity of the bright line reflexion spectrum. In absorption, on the other hand, this is of less importance, for the cloudiness in question does not help the transmission of light which is selectively reflected by the structure of parallel planes: thus the blackness of the absorption lines is not prejudiced. To study the absorption spectrum, a plate was cut from the shell already used in making the transverse section of fig. II. The plate was of course parallel to the strata, and was polished on both sides. Its thickness was about 0·15 mm. The transmission spectrum is shown in Plate 8, fig. III. Sky light was used as the source, and the Fraunhofer lines serve for scaling the spectrum. Two sinuous dark bands cross the spectrum, one of these is in the yellow and the other in the blue region. When the former coincides with the D lines, the latter very nearly coincides with G. The corresponding sinuosities of the two bands show without doubt that they are due to the same structure, the same irregularities of linear period affecting each of them. It is clear also that the band in the yellow cannot result from a first-order reflexion, for in that case the next order would be the second, and this would give an absorption in the ultra-violet at $\lambda 2950$, thus at the extreme end of the solar spectrum. In fact, however, the next absorption band appears at about $\lambda 4300$. The conclusion must be that the yellow band results from a second-order, and the blue from a third-order reflexion. This agrees with the observed position of the bands as well as can be expected. They are too broad and diffuse to be located very closely in the spectrum. The diffuseness of the bands corresponds to the raggedness of the stratified structure (Plate 8, fig. II). This is in striking contrast with the structure of chlorate crystals shown in the photographs accompanying the previous paper.

1.



2.





3.

The conclusion that the chief visual reflexion is of the second-order, is the same as that reached by Pfund (*loc. cit.*) who located the first-order in the infra-red.

Taking the refractive index of the material, whether calcite or horny matter, as 1·6, the linear distance between successive reflecting planes for second-order reflexion should be 0·00037 mm. If the alternate strata are of different material, this makes the complete linear period 0·00074 mm. Magnified up to the scale of fig. II, Plate 8, this becomes 0·296 mm., or 3·4 complete periods to the millimetre. This seems to agree with the period on the photograph, so far as the latter admits of measurement. The curious blank spaces where the structure seems to be in abeyance make this rather difficult. It would seem probable that these blank spaces consist of horny matter without calcite.*

§ 4. *Summary.*

The structure giving rise to the colour of mother-of-pearl is investigated. The results generally are confirmatory of those given by Brewster and A. H. Pfund. Micro-photographs are reproduced showing the grating structure of a pearl-oyster shell, and the structure of parallel layers of an "ear" shell. The absorption spectrum of the latter is also reproduced. In agreement with the observed spacing of the layers under the microscope, it shows the reflexion to be of the second order.

DESCRIPTION OF PLATE 8.

1. Grating structure of pearl oyster, section parallel to plane of shell. × 80 diameters.
2. Transverse section of ear shell showing stratified structure which gives rise to second-order coloured reflexion. × 400 diameters.
3. Transmission spectrum of a section of the same ear shell cut parallel to the strata. Shows irregular absorption bands in the yellow and blue.

---

* I cannot claim to have closely studied the zoologicall iterature on the minute structure of shells.

Investigations on the Paramagnetic Sulphates at Low Temperatures.

By L. C. Jackson, M.Sc., A.Inst.P., "1851 Exhibition Research Scholar," and
Prof. H. Kamerlingh Onnes, For.Mem.R.S.

(Received January 12, 1923.)

The investigations on paramagnetic substances at low temperatures, already carried out in the Physical Laboratory of the University of Leiden,* have led to two results of special importance, viz.: (1) that while practically all such substances obeyed the law of Curie in the generalised form of Weiss, which we write $\chi(T+\Delta) = \text{const.}$, down to about 80° Kelvin, some of them showed deviations from this law at lower temperatures, such that the susceptibility χ increased *less rapidly* with fall in temperature T than is given by the above law; and (2) that in a series of substances of similar constitution, but containing different amounts of water of crystallisation, the constant Δ in the above expression was smaller the greater the amount of water of crystallisation present, and that substances, the "inert part" of the molecule of which was large compared with the "paramagnetic part," as in ferric alum, possessed very small values of Δ. The value of Δ was thus *smaller* the *greater* the *magnetic dilution* of the molecule, and hence the greater the distance apart of the paramagnetic atoms of the molecule.

The present paper is intended as a preliminary communication, summarising the results of further work undertaken to corroborate and extend the above-mentioned observations. It is hoped to publish a full account of the work shortly.

Measurements have been carried out with the following series of compounds of increasing "magnetic dilution":—The anhydrous sulphates, the heptahydrated sulphates, and the ammonium double sulphates of cobalt and nickel. From the results obtained it became immediately obvious that other types of deviation from the Curie law than that mentioned above can occur at low temperatures. Thus, cobalt ammonium sulphate and nickel ammonium sulphate show deviations from this law, such that the susceptibility increases *more rapidly* with fall in temperature than is given by the law; hence the graph of $1/\chi$ against the absolute temperature is a straight line for temperatures above about 80° K., but becomes concave towards the temperature axis at lower temperatures. Further, a maximum and minimum of susceptibility may·occur in the region of liquid hydrogen temperatures (14°–20° K.) (anhydrous nickel sulphate), and again the $1/\chi.T$ curve may have an inflection at low temperatures (cobalt sulphate heptahydrate and nickel sulphate heptahydrate), thus showing deviations from the law in opposite senses in different parts of the curve.

---

* 'Leiden Comm.,' Nos. 122*a*, 124*a*, 129*b*, 132*e*, 139*e*.

Secondly, the dependence of the Δ's on the content of water of crystallisation, or more generally on the "magnetic dilution" of the substance, is not so simple as was thought at first from the previous results cited above. Thus, for the series of nickel compounds the Δ's have the following values :— $NiSO_4$, $+79{\cdot}4$, $NiSO_4.7H_2O$, $-62{\cdot}2$, $NiSO_4(NH_4)_2SO_4.6H_2O$, $+4$, and for the cobalt compounds $CoSO_4.7H_2O$, $+13{\cdot}7$, $CoSO_4.(NH_4)_2SO_4.6H_2O$. $+22$. The values of Δ in a series of similar substances do not follow necessarily the order of the "dilution" of the substances, and may be positive or negative in the same series of compounds.

It will be seen that the later results show that the phenomena exhibited by paramagnetic substances at low temperatures are more complex than was indicated by the earlier work. The existence of the type of deviation from the law $\chi(T+\Delta) = \text{const.}$ shown by cobalt ammonium sulphate has also been proved by Foëx* for the case of one of the principal susceptibilities of siderose at fairly low temperatures.

Many attempts have been made to find a theoretical explanation of the behaviour of paramagnetic substances at low temperatures. Every investigator, with the exception of Foëx,† has however considered only the possibility of the existence of the type of deviation found in the earlier work. The later results given above receive no explanation on any of the present theories except that of Foëx. The latter can give at least a qualitative explanation of all the phenomena described above, by taking into account in the calculation the effect of the crystalline structure of the substances on the behaviour of the elementary magnets.

---

* 'Thèse, Strasbourg,' 1921.

† *Loc. cit.*, chap. 6.

Investigations on the Paramagnetism of Crystals at Low Temperatures.

By L. C. JACKSON, M.Sc., A.Inst.P., "1851 Exhibition Research Scholar," and
Prof. H. KAMERLINGH ONNES, For.Mem.R.S.

(Received, January 12, 1923.)

Although the investigation of the principal susceptibilities of paramagnetic crystals especially at low temperatures promises results of very great interest both from the theoretical and experimental standpoints, data on the subject are almost entirely wanting, probably at least in part as a result of the difficulty of making the necessary observations. Thus the only record of crystal susceptibilities as functions of the temperature is that of the measurement of the principal susceptibilities of siderose by Foëx,* but these observations were not extended to lower temperatures than those obtainable with liquid air.

The present preliminary note gives the results of an investigation of the principal susceptibilities of two crystals, namely, cobalt ammonium sulphate and nickel sulphate heptahydrate as functions of the temperature down to about 14° K. (melting point of hydrogen). These are thus the first observations of this type in the region of very low temperatures. The substances were chosen as representative of different classes of crystals (cobalt ammonium sulphate is monoclinic and nickel sulphate heptahydrate is rhombic), and as they promised results of special interest in the light of the phenomena shown by the mean susceptibilities obtained from measurements on the substances in powder form (see foregoing paper).

The measurements with cobalt ammonium sulphate showed that this substance follows the law of Weiss $\chi(T+\Delta) = C$ down to about 80° K. for each of the three principal susceptibilities, and that the constant C is the *same in the three cases,* thus confirming a similar observation by Foëx for the case of siderose. The greatest susceptibility brings with it the smallest Δ and *vice-versâ*: hence the magnitudes of the deviations from the simple Curie law $\chi T = C$ are in the inverse order of the magnitudes of the principal susceptibilities.

Further, each of the principal susceptibilities deviate from the law $\chi(T+\Delta) = C$ at temperatures obtainable with liquid hydrogen, and the deviations are all in the direction such that the susceptibilities increase *more rapidly* with fall in temperature than is given by the above law. The magnitudes of the deviations are different in the three cases, being greatest for the axis of smallest susceptibility and *vice versâ*.

---

* 'Thèse, Strasbourg' (1921), p. 81.

Nickel sulphate heptahydrate shows similar phenomena, the principal susceptibilities following the law $\chi(T+\Delta) = C$ down to about 100° K., and the value of C being equal in the three cases. At lower temperatures each of the principal susceptibilities shows deviations from the generalised Curie law of the same type, as shown by the mean susceptibility (see foregoing paper).

Another point of considerable interest may be noted. It was suggested, as a result of the earlier measurements on powdered substances, that the value of Δ depends in a simple way on the "dilution" of the substance, and hence on the distance apart of the paramagnetic atoms, and also that in crystals the Δ's for the three principal susceptibilities might be directly dependent on the spacing of the paramagnetic atoms along the three principal magnetic axes. It has now been shown that the first of these suggestions is not altogether justified. It remained to test the second. If this one is right, a crystal of gadolinium ethyl sulphate having the same dilution as gadolinium sulphate, for which $\Delta = 0$ was found in the powder though crystallographically monoclinic, ought to be magnetically isotropic. This proved to be the case with beautiful samples we owe to the kindness of Prof. Urbain, of Paris. The suggestion was further tested in the following manner, which is avowedly of a provisional nature. Taking into account the known crystallographic data an attempt was made to arrange the nickel atoms in nickel sulphate heptahydrate on a space-lattice in such a manner that the Δ's were a *linear function* of the spacings of the nickel atoms in the corresponding directions. It was found that this could be done, thus providing a hypothetical space-lattice. Prof. Keesom, of Utrecht, then had the goodness to test this hypothetical space-lattice by submitting the substance to X-ray analysis by the Debye-Scherrer method. He now reports that, provisionally, the actual arrangements of the nickel atoms and the relative dimensions of the space-lattice agree with the hypothetical one suggested from the magnetic measurements. It has not yet been found possible to test the hypothesis of the linear relation between the Δ's and the spacings in the case of cobalt ammonium sulphate.

It may thus be stated that, should the hypothesis be confirmed by further investigations, the measurement of the principal susceptibilities as functions of the temperature may serve to determine the structure of paramagnetic crystals, at least as concerns the positions of the paramagnetic atoms.

Pleochroic Haloes of Various Geological Ages.

By J. Joly, Sc.D., F.T.C.D., F.R.S.

(Received November 24, 1922.)

[Plate 9.]

Introduction.

This paper is concerned with the radio-active phenomena apparent in micas of various Geological Ages, but more especially in the black micas of Ytterby and Arendal. They are, I believe, sufficiently remarkable to merit record, even if much that has been observed remains but partially explained, or even not explained at all.

The subject of "reversal" enters conspicuously into a consideration of the observations on the more ancient haloes. The theory of reversal (using the word in the photographic sense of the term) originated in the difficulty of finding in any other way an explanation of the sequence in which the structural features of haloes develop.* The haloes then under consideration were the Devonian haloes of the Co. Carlow (Ballyellen) mica. On the reversal theory these haloes would be described as "positives." In Archæan micas, under conditions of longevity far in excess of the palæozoic haloes, developments have occurred which can be brought under the same theory; but the haloes concerned must be described as "negatives." Whether in these very ancient micas we have positives or negatives seems to turn on the radio-active strength of the nucleus. Again we have Uranium haloes in which some parts are negative (reversed) and others positive.

Furthermore, certain other radio-active phenomena in the mica show positive and negative developments side by side.

The most interesting outcome of these supposed reversal effects appears in the existence in the Ytterby mica of certain very rare and minute objects—quite colourless and spherical in form. If radio-active in origin the nuclear particle placed centrally in these haloes must emit Alpha rays of very low velocity—possessing in fact a range of no more than 1·5 cms. in air. The element concerned is at present unknown, or at least its radio-active properties have not been recognised. It is unfortunate that these minute haloes seem restricted to rare groups or clusters, the scarcity of which is shown by the fact that protracted search in some score of samples of the mica has failed to reveal even a single group comparable with that first found; although a very few individual haloes have been met, but lacking the definition of the original group. It is probable that the primary nuclear conditions of very minute dimensions and strong radio-active charge are seldom combined. Dark or positive developments of these haloes have not been definitely observed. It is possible that they would be difficult to detect in the obscurity of this mica. However, minute positive rings of such dimensions as would seem to ally them with the small haloes have been detected.

---

* 'Phil. Trans.,' A, vol. 217, p. 57.

An account of the little that has been ascertained respecting these haloes occupies the second part of this paper.

Following on this a section of the paper is devoted to description and illustration of haloes smaller than Uranium haloes and possessing structural features which clearly separate them from connection with this family of radio-active elements. The possibility of reconciling them with Thorium haloes, assumed to be severely metamorphosed, is discussed. I do not believe that the view that they are so derived will be accepted. Meanwhile I have contented myself by calling then X haloes. They would seem to reveal the existence of yet another unknown radio-active element.

The last part of the paper is devoted to comparative observations on the dimensions of incipient Uranium haloes—first-ring haloes—occurring in micas of very different antiquity. If confirmed by further investigation the results would, I think, possess interest quite apart from the subject of the estimation of geological time by the Uranium-lead ratio. For the most direct explanation of the results—the exhaustion of a former, relatively short-lived, isotope of Uranium—may be of considerable significance and seems to throw a suggestive light on the history of the elements over great time-epochs.

Bleaching of the Archæan Micas.

A striking feature of the black Archæan micas of Ytterby and Arendal is the presence of perfectly colourless or bleached areas, which in all cases seem associated with radio-activity.

In some samples of these micas the most conspicuous instances of bleaching are found in bands along which minute particles are centrally distributed. These particles may be opaque or translucent. They are sometimes arranged with remarkable precision along the axis of the band. At other times they occur in small groups; each group being attended with a widening of the bleached area. The photographs, Plate 9 (figs. 1 and 2), show bleaching around such particles. It is to be noted that areas where particles appear to be absent by no means indicate the real absence of such particles. Bleaching probably *never* occurs without the inciting cause being segregated in mineral particles. In some cases, however, some of these may be removed in the adjacent cleavage mica flake. Examination of the photographs will show that the bleaching is attended, in many cases, by a border of darkened mica. In general this external border is very conspicuous. This border being undoubtedly radio-active in origin, and no inciting cause for its existence, other than the central particles, being forthcoming, there is little doubt that the bleaching is also a radio-active effect. In short the inference is irresistible that the central colourless area is referable to the reversal of the more intensely affected region bordering on the centrally placed particles.

The colourless bands take on extraordinary forms, sometimes branching and crossing irregularly and in any direction. In other cases they follow a straight course of wonderful regularity. The width of such bands appears to depend on the distribution of the central particles. There is some indication, however, that the radio-active influence of these particles is not always equally penetrating. A few bands have been measured which exhibited a width of no more than about 0·01 mm.; the radial effect of the very minute central particles in this case extending no further, apparently, than about half this distance. This observation would appear to suggest that the responsible radio-active substance might be the same as that which originated the minute haloes to be described in the ensuing section of this paper.

In certain samples of mica, bleaching is absent or rare. Heavy staining along linear lines and bands may be present, such as are not uncommon in younger micas.

The radio-active origin of bleaching is most convincingly displayed in the case of certain haloes of known radio-active origin. The drawing (fig. 3) made from micrometric measurements (to a scale of 1000 diams.) of a reversed Ytterby Uranium halo of rather exceptional beauty, plainly reveals what has happened. In juxtaposition to it I have shown a half-halo to the average dimensions of a Ballyellen (Devonian) halo ('Phil. Trans.,' *loc. cit.*). It is apparent that the dark bands of the normal halo are replaced by bleached bands. Such Archæan haloes may show only one of the bands reversed or the central area (as in fig. 2, Plate 9); or, again, the central area may be dark and the external features of the halo strongly affected by bleaching.

It is interesting to note that reversal in these cases may bleach not only the darkening we would normally ascribe to radio-active effect, but also the colour which generally prevails in the mica. I have endeavoured to indicate this in the drawing. In some cases this complete and radical bleaching of the mica

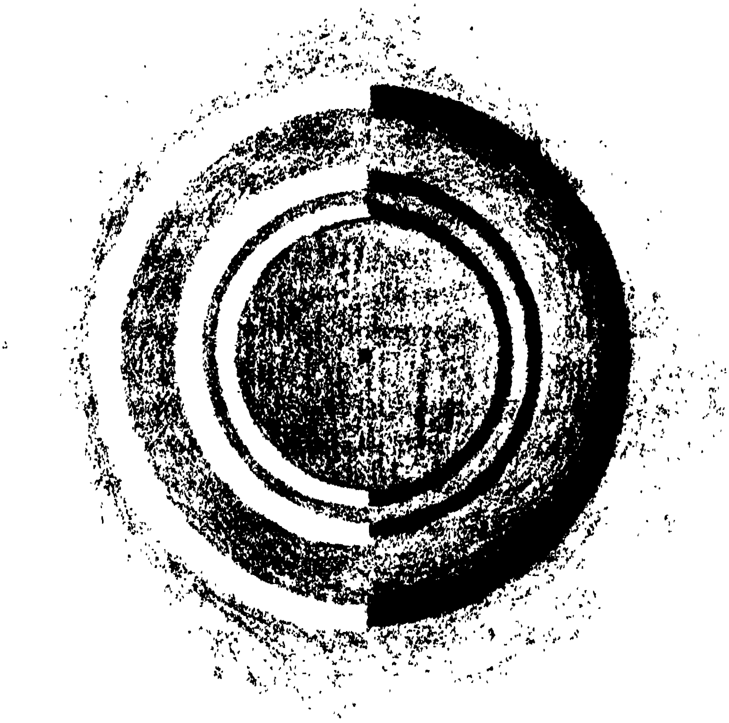


FIG. 3.

is not observed; the bleaching area is then simply reduced to the colour of the surrounding mica. The photograph (Plate 9, fig. 6) of an X halo, to be referred to later, shows a single band of complete bleaching, as will be evident on comparing the bleached area with the comparative obscurity of the containing mica.

Haloes in comparatively early stages of development may show bleaching. Thus, small, quite colourless haloes may, occasionally, be met with which are, apparently, due to radium emanation. In by far the larger number of cases, the earlier forms are in deep sepia tints or merging to black. In short we may find, along with bleached or partially bleached haloes, others which are in all respects positives and similar to haloes of younger geological age, save in so far as they partake of the more or less turbid character of the ancient micas.

It is difficult to ascribe the bleaching to any purely chemical effect. As already stated, it is nearly always definitely associated with a nucleus and very often with characteristic radio-active staining. At one period I was led to believe that these bleaching effects were accelerated by heating the mica ('Nature,' April 22, 1922). A preliminary experiment on the Ballyellen mica gave results of a nature strongly to support this view, but further careful work by Dr. J. H. J. Poole has failed to substantiate it. Using a micro-electric furnace, enabling a halo to be kept in observation while the temperature was gradually raised, no case of bleaching was observed. At a very low red heat the halo and the surrounding biotite alike darkened up, till almost complete opacity was obtained and all signs of radio-active effects disappeared. There may, however, be some condition which obtained in the preliminary experiment and which was absent from subsequent ones. Possibly the stage of development of the halo may be important. The matter is deserving of further investigation.

So far as our knowledge of the subject goes, I think the reversal theory best explains the observations. Recurrent reversal in photography is well known. In this case the character of the developed image varies from positive to negative periodically under prolonged exposure. It is quite possible that a similar effect may take place in very ancient haloes having strongly radio-active nuclei. The somewhat capricious distribution of bleached and dark haloes, or even the distribution of bleaching within any one halo, may be ascribable to this phenomenon. Experimental production of reversal in biotites, using strong radio-active charges, might be possible. However, the condition of long-continued radiation may be essential.

Hibernium Haloes.

Some years ago I made a preparation of cleavage flakes from a specimen of black Ytterby mica taken from the Museum of Mineralogy and Geology, of Trinity College. In one of these flakes a cluster of very minute colourless and spherical objects was observed. These same objects were found in some other of the flakes, but not in all. Subsequently other preparations from this sample of mica were made, and from various exterior parts of it. No more of the small objects were found. The sample itself appears to have been lost during the occupation of my laboratory by troops in Easter week, 1916; the week of the Sinn Fein rebellion in Dublin.

Examination of the colourless objects showed that in every case a central nucleus was present as in radio-active haloes. But haloes previously studied were invariably depicted in greater depth of colour than the surrounding mica. The possibility that loss of colour might have originated from a phenomenon similar to photographic reversal or solarisation had not then occurred to me. The reversal theory, as affording an explanation of the order of development of certain halo features, was thought of only at a later date ('Phil. Trans.,' *loc. cit.*, p. 57). Subsequently I found, as stated above, haloes showing all the appearances of reversal in this Ytterby mica and in the mica of Arendal. Similar appearances characterise the linear areas of radio-active staining. Again, certain reversed incipient halo-structures appear to possess much the same characters as are shown by the minute spheres under consideration. In the light of these observations, I can come to no other conclusion than that the minute bleached spheres have originated in just the same manner as bleached Uranium haloes.

No coloured representatives of the minute spheres have been found, although a curious oolitic or roe-like appearance in neighbouring parts of this mica, showing in very faded tints, suggests the possible obliteration of such minute haloes. It must be borne in mind that this Ytterby mica has been exposed to severe metamorphic actions, in the course of which those foreign elements were introduced into it which have made it famous in the history of chemical science. This appears in the analyses quoted in another section of this paper.

In figs. 4 and 5, Plate 9, photographs of a field of these haloes are reproduced; a small part of the great cluster referred to. They give only an inadequate idea of the appearance presented in the microscope.

It must be realised that these haloes are of diametral dimensions less than the thickness of the containing cleavage flake; in this respect reversing the condition under which larger haloes are usually examined. A measurement of the thickness of the mica flake showed that it was from 0·018 to 0·020 mm. The haloes are about 0·010 mm. in diameter. As their centres are located at various depths in the mica it is evident that only a certain number of them can exhibit their full radial dimensions. It follows, too, that the definition will vary with the amount of the halo-sphere contained in the flake. In the photographs a well defined boundary can only be expected of haloes which are not only completely contained in the flake, but are correctly placed in the external focal plane of the object lens. On these accounts higher magnification was not found to yield any better photographic results.

A very remarkable feature of these halo-spheres is the brilliantly luminous appearance presented by the nuclei. They stand out with all the appearance of actually emitting light. At one time I held the view that these luminous points were not, in fact, nuclei but focal points determined by the refracting properties of the sphere. This theory fails in view of the phenomena attending multiple nuclei next to be described; but that a raised refractive index characterises the material of the halo-spheres seems a necessary inference in order to account for the luminosity of the nuclei.

The photographs show that exactly in the centre of some of the haloes is a very small particle. For such particles to be reproduced in the photograph the focus must be very perfect. Accordingly only in the case of a few of the haloes is the nucleus sharply defined. In some cases more than one nucleus will be discernible. If there are two such particles separated a little the halo takes on an oval or ellipsoidal form, the major axis connecting the particles. Further separation of the particles produces a dumb-bell shaped halo. Three particles, if located at the apices of an isosceles triangle, give rise to a pear-shaped halo. A group of such nuclei give simply an enlarged halo-sphere, and so on. The formation of the halo being plainly dependent on the disposition of the nuclei, and this influence extending in each case a fixed and definite distance.

The central particle or nucleus appears, in every case, to be composed of a highly refracting, transparent and colourless substance. No crystalline form has been observed. They are very minute and under certain conditions of focus may assume an appearance of blackness and opacity, due, probably, to the optical properties of the containing halo-sphere. Careful examination, using an oil immersion, leaves, however, no room for doubt that they are in fact limpid and colourless.

The majority of these haloes possess dimensions in sufficient agreement to strongly support their radio-active origin. Thus on 23 of these haloes, having each a single nucleus, the micrometer head read as follows—the magnification being such that 54·4 scale divisions = 0·01 mm.:—

56, 60, 57, 57, 52, 59, 52, 58, 55, 50, 55, 56, 53, 53, 52, 57, 54, 50, 55, 55, 49, 57, 54. Mean, 54·6. Hence the radius is 0·0050 mm.

Subsequently 50 of these objects were measured, using a lower magnification, *i.e.* 20 scale division = 0·01 mm. The extreme readings were 18·5 and 23·5, the mean being 20·8. This gives $r = 0\cdot0052$. This neglects the nuclear deduction which would be only a small fraction of the radius. It is to be noted that the 50 haloes dealt with in the second observation doubtless included many which enter into the first observation.

But there seem to be present a small proportion of these colourless objects larger in dimensions, the greater size of which cannot be referred to multiple nuclei. Thus two such haloes read 34·0 and 34·8 scale divisions to the magnification used in the last series of measurements. This gives $r = 0\cdot0086$. Several haloes were found to possess dimensions sufficiently close to this to entitle them to be regarded as originating by rays of like penetration. The existence of such larger and smaller haloes side by side is paralleled by the intermingling of more fully developed haloes with swarms of first-ring Uranium haloes in the Ballyellen mica. It is possible that a yet larger size exists, but the evidence for this is, on the available material, hardly adequate.

The difficulty may be raised that if these haloes are all alike reversed haloes we would not expect reversal of haloes of different stages of development to prevail; those of fuller development should, in fact, be more prone to reversal than those in an earlier stage. However, the reversal of immature haloes as well as of fully developed haloes in the Ytterby and Arendal micas has been observed, as already stated.

As regards the mode of origin of these remarkable objects, the case for a radio-active origin rests upon the following considerations:—

(1) The agreement in radial dimensions of a large number is just as striking as in the case of halo-clusters generally.

(2) The development around the nucleus is concentric to a high degree of accuracy.

(3) The relation of the external form to the distribution of nuclei, when more than one are present, is such as to show that each nucleus affects the surrounding mica to an equal radial distance.

(4) The colourlessness of these objects is no argument against a radio-active origin, as follows from the instances of reversal or bleaching around nuclei which have generated halo-structures of dimensions characteristic of known, recognised haloes.

(5) The dimensions of the nuclei appear to be too minute to admit of the view that they have, by chemical effects, bleached the surrounding mica. The diametral dimensions of nucleus and halo-sphere may be in the ratio of 2 : 50 or 3 : 50 in scale divisions of the micrometer. The relatives volumes are, therefore, as 1 : 15,600 or as 1 : 5000, and there is no appearance of disintegration or loss associated with the nucleus. It remains clear and sharp; isolated in limpid mica.

(6) The outside dimensions of the halo may be sensibly alike when the nucleus varies appreciably in size. This is in favour of a radio-active origin and not of a chemical origin.

(7) These objects are spherical. The evidence for this is the way they affect the transmission of light through the mica. Now it is not to

be expected that chemical effects would be transmitted equally along and across the cleavage. Inclusions in mica, *e.g.* hæmatite, invariably are spread in the plane of the cleavage. Radio-active effects are not affected by the cleavage. Serial cleavage sections of haloes demonstrate this fact. A bleaching effect due to a chemical influence spreading in the cleavage plane would effect the opacity of the mica in a relatively feeble degree and would not be limited by such a gradation of opacity as is observed surrounding the haloes.

(8) If bleaching is due to chemical effects, how are the phenomena of bleaching treated in foregoing pages to be explained, *i.e.* the reversal of halo rings of recognised characteristic dimensions ?

It is of interest to note that minute halo rings, developed in dark sepia or smoky tints, are occasionally found in the Ytterby and Arendal micas. They are of perfect circularity, and only found where a very small nucleus has operated to produce some larger halo structure. This structure may be part of a halo of undoubted radio-active origin. For instance it may be an X halo as described in this paper. It may be a Radium emanation halo. It may be a normal Uranium halo in some early stage of development.

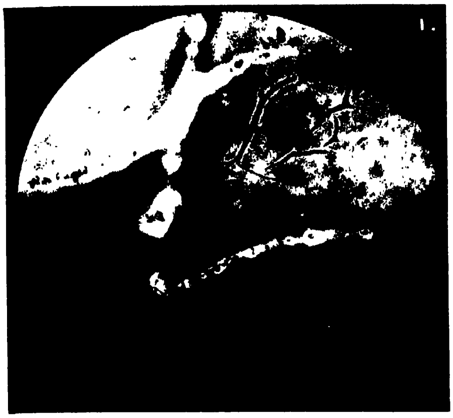
Are these rings, which are very scarce, optical effects only ? I find it difficult to account for them as such. To see them the lighting must be excellent and the central nucleus in sharp focus. I have endeavoured to photograph them but with only partial success. The focus difficulties arising upon projection are great.

Mr. L. B. Smyth and I have read these minute rings independently and arrived at almost identical dimensions. They are fairly sharp objects. Thus one of those rings which I read as 0·0051 was independently read by Mr. Smyth as 0·0050, external radius.

The rare occurrence of such rings within Uranium halo-structures is, I think, evidence that the responsible radio-active element—if their origin is radio-active—is not genetically connected with the Uranium family of elements, but is only accidentally associated with it. Whether they are connected with the small Ytterby bleached halo-spheres is a surmise at present incapable of proof.

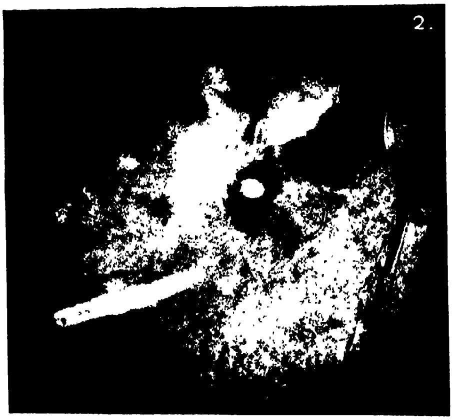
It is of interest to consider the air-range of an Alpha-ray competent to generate one of these small spheres. The radial dimension given above is probably not a full measure of the entire range in air. This is so because the Bragg-Geiger curve does not generally depict the halo out to the full range. The crest of the curve is about 0·6 cm. short of the extreme range. The

Fig. 1.



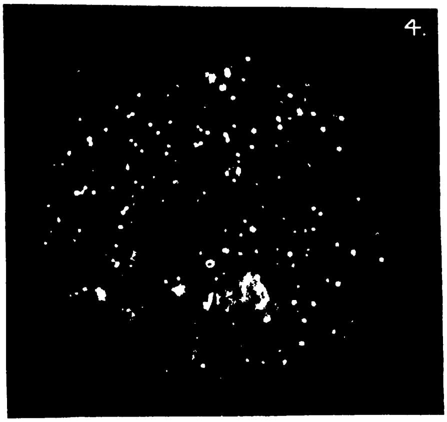
× 110 diams.

Fig. 2.



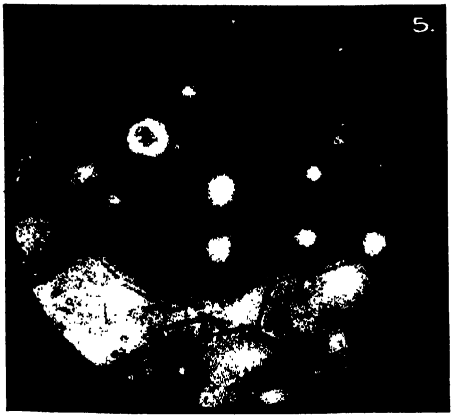
× 110 diams.

Fig. 4.



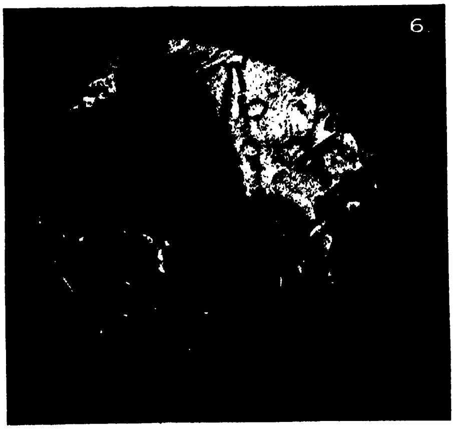
× 100 diams.

Fig. 5.



× 400 diams.

Fig. 6.



× 256 diams.

effective limit of the curve may be—as measured in air—some three or four millimetres behind the extreme range.

If the radius of the small haloes is 0·005 mm. and the equivalent of 1 cm. range in air for this mica be 0·000436 cm., the air equivalent of the halo-radius is 1·15 cms. If to this we add 3·5 mms. to arrive at the extreme range of the Alpha-ray, we find that these haloes involve a range of 1·5 cms. in air. The haloes of larger radius would arise from an Alpha-ray possessing a range of about 2·25 cms. in air. The shortest range at present recorded is that of the Alpha-ray of U_1, *i.e.* 2·50 cms.

The range 1·5 cms. in air is also that to which the generating ray of the small dark rings (*ante*) must be referred. It is certain that the responsible element cannot be placed upon the Geiger-Nutall curve for the Uranium family. Extrapolation on this line to a range of 1·5 cms. shows that the rate of decay of the unknown element must be some 2000 millions of times slower than that of Uranium. Basing his conclusion upon my first evaluation of the range of the ray responsible for the generation of the small haloes, *i.e.* 1 cm. in air, Mr. Svein Rosseland, of Copenhagen, arrived from theoretical principles at the conclusion that the element in question might be Yttrium ('Nature,' June 3, 1922). It may turn out that this is the case. Meanwhile, I have suggested the provisional name "Hibernium" for the unknown and, at present, hypothetical element.

The X Haloes of Ytterby and Arendal.

In the black Archæan mica of Ytterby and Arendal are certain haloes which call for special consideration. A photograph of one of these haloes, to a magnification of 256 diameters, is appended (fig. 6, Plate 9). Till more is ascertained respecting these haloes I propose to call them X haloes. Their radio-active origin is beyond question.

A feature of X haloes—which, however, seems to be sometimes absent—is the presence of a bleached (reversed) belt or ring. Very often this ring is beautifully sharp and definite, and as the bleaching of the mica is complete—its own natural colour being destroyed—the effect is very striking, especially when, as sometimes happens, these haloes are so closely congregated as to overlap.

It is evident that the photograph represents a strongly developed halo. Its structural details are not fully revealed in the photograph.

X haloes may vary somewhat in dimensions. These variations may be those common to all haloes, *i.e.* such as depend upon the stage of development.

But there appears to be another source of variation. The bleaching of a portion of the halo may obliterate certain structures and may even create deceptive structural appearances. In figs. 7 and 8 I give drawings (to a scale of 1000 diameters) of two types of X halo, which at first I had regarded as distinct and of different radio-active origin, but which I now believe to be variants of a common type and origin.

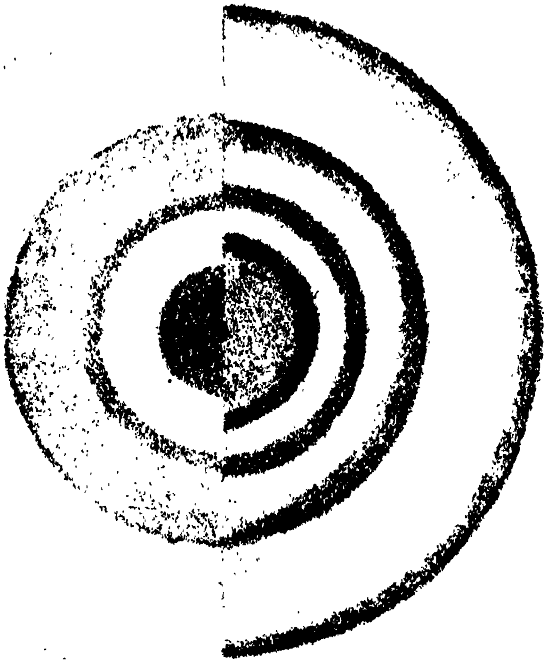


Fig. 7.

On account of a certain degree of resemblance to the Thorium halo I have drawn a half-Thorium halo in juxtaposition to a half-X halo in fig. 7. The X halo shown in this figure is approximately of the dimensions of the halo reproduced in the photograph (fig. 6, Plate 9). The outside radius is 0·0280 mm. A wide, uniformly shaded band is the first outside feature. This is limited inside by a fairly strong dark band having an outside radius of 0·0170. Within this band bleaching begins at a radial distance from the centre of 0·0152. The bleaching extends inwards till the central feature is met. This is a deeply coloured disc with a radius of nearly 0·009 mm.

But these structural features may vary considerably. Thus the outside radius may be as little as 0·0274 mm. and as much as 0·0294. The external radius of the dark ring may vary from 0·0160 to 0·0188. The external radius of the bleached band may vary from 0·0146 to 0·0170, and the central disc from 0·0074 to 0·0098. These are the limiting readings obtained from a considerable number of such haloes.

The type of X halo shown in fig. 8, and designated X_2, at first sight looks very different. But it must be remembered that the absence of a reversed band is not really a distinctive feature. Reversal, if indeed it is such, is only a stage of development and probably a recurrent stage. This second type of X halo is specially abundant in the Ytterby mica. In that specimen of this mica which contains the Hibernium haloes they are very numerous. Here they present a faded appearance, their colour and intensity approximating to the surrounding mica, so that they might easily be overlooked on hasty examination. Their dimensions are, however, quite determinate. The external radius varies around 0·028 mm. It has been read from 0·0265 to 0·0292. The uniformly shaded band which surrounds the central disc varies in outside radius from 0·0133 to 0·0157. The central disc has been read from 0·006 to 0·0084. The lower value was obtained from a rather obscure specimen. In fig. 8 I have shown the halo features as reading 0·0280. 0·0144, 0·0080 successively.

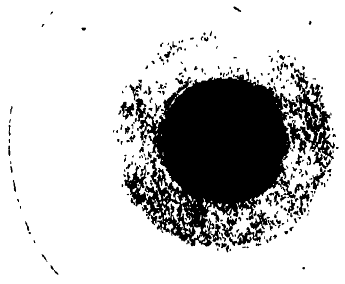


Fig. 8.

It may be asked upon what grounds are these two types of haloes regarded as alike in origin? The answer is that it is not improbable that the band surrounding the central disc of fig. 8 is capable of outward development along a curve of ionisation which slopes steeply outwards. If now we assume that it develops to the radial limits of the dark ring encircling the bleached area of fig. 7, and that reversal obliterates the colouration from the central disc outwards, developing a bleached band somewhat similar to that shown in fig. 7, then the two haloes would be identical. This correlation is not impossible, nor even improbable. To accept it is, I think, preferable to the assumption that these haloes are of distinct radio-active origin.

Considering the successive features of X haloes in comparison with Thorium haloes there are certain features in fair agreement. In the drawing of the half-Thorium halo I have depicted the average radial limit of the "pupil" by the third ring from the centre. In point of fact when the pupil develops it is most often found that the inner features become blackened up and undistinguishable. The radius of the Thorium pupil is a somewhat variable dimension. It has been measured from 0·023 to 0·028 ('Phil. Trans.,' *loc. cit.*, p. 70). It is, in short, reconcilable with the outside boundary of the X halo. The second Thorium ring from the centre is again not in serious disagreement with the dark band, which in the X halo may encircle the bleached area. If, then, some metamorphic effect had obliterated a former "penumbra" appertaining to the X halo there would be a certain amount of resemblance between the Thorium halo and an X halo "restored" according to the foregoing assumption.

On the other hand it is to be remarked that in some scores of X haloes no trace of any outside feature has been detected. Nor has any trace of the first Thorium ring been observed, even when there was no bleaching. This ring is, in most cases, a very distinctive feature of the Thorium halo, corresponding to the first crest of its curve of ionisation. Again, Thorium haloes exist in these micas with the usual features, and often in the close neighbourhood of the X halo. What secured these normal Thorium haloes from the assumed destructive effects? I do not think the view that X haloes are metamorphosed Thorium haloes is at all probable.

The Dimensions of Primary Ring Uranium Haloes in Micas of Various Ages.

In dealing with the genesis of pleochroic haloes ('Phil. Trans.,' A, vol. 217, pp. 51 *et seq.*) attention was drawn to the fact that the curve which expresses the added effects of the ionisation due to the several Alpha-rays emitted in the radio-active transformation of the Uranium series, exhibits a single maximum of ionisation at a distance from the axis of Y, which is 2·2 cms. in air, but which in the haloes of Ballyellen (Co. Carlow) is found to occur at a distance of closely 0·0142 mm., this latter distance being about 12 per cent. greater than is deducible from the air-curve as modified for the stopping properties of the mica.

The stopping power, according to a law deduced by Bragg and Kleeman, depends upon the chemical composition and the density of the medium. This law, as applied to the mica, is confirmed by the fact that the other features of Uranium haloes, the features of emanation haloes, and those of Thorium haloes, are found to be in general agreement with the corresponding air-curves as modified according to the law referred to.

As no reconciliation of the discrepancy is immediately apparent, the suggestion was offered that the range of one or more of the primary short range Alpha-rays which attend the first stages of radio-active change of Uranium had diminished during geological time; an effect which might be due to the former existence of a proportion of Uranium atoms having properties other than those which obtain in the Uranium of our time. It was also suggested that such a change in range might be revealed if the dimensions of haloes of very different geological ages could be compared.

At the time when this was written I had failed to find available haloes in recent micas. My search was more especially directed to the finding of the initial halo structures in the Tertiary mica of the Mourne granite. It was, therefore, with much interest that early in this present year (1922) I found in the cleavage flakes of the porphyritic mica developed in the druses of the Mourne granite a few of these primary halo rings. It is remarkable that subsequent research, although it added several haloes available for measurement, afforded nothing so definite as this early find. The measurements to be cited presently are mainly readings made on the cluster of delicate halo rings as first found. Other rings are also dealt with.

The Mourne mica is not deeply coloured. Generally it exhibits yellow or blue tints, which tints are in many cases disposed in wide bands around the margins of the flakes, the central areas being but faintly coloured. The halo is depicted in brown or smoke-coloured tints. Its dimensions are not found to vary with the colour of the containing mica. That it is developed from the Uranium radio-active series is shown: (*a*) by the approximate agreement of its radial dimensions with the corresponding halo rings of the Ballyellen mica; (*b*) the occasional existence in near association of characteristic radium emanation haloes. In these the central conduit may be sometimes detected (as a perforation in the mica) as well as the outer ring due to RaC; (*c*) no known haloes originate with a primary ring of similar dimension. It may be added that Thorium haloes seen to be absent from this mica. The absence of fully developed Uranium haloes must be ascribed to the comparatively recent origin of this granite. Although large inclusines with heavy brown staining around them are not uncommon, nuclei sufficiently small to generate well defined haloes have not had time to give rise to colour effects in the outer parts of the halo-structure.

It is necessary to investigate the stopping power of this mica. An analysis by Haughton is quoted in Teall's 'British Petrology' (p. 229). In order to verify this more directly with regard to the particular mica containing the haloes, a sample of this mica was analysed by Mr. Stone, of the Chemical Department of the Royal Dublin Society. The two analyses are as follows:—

	Haughton.	Stone.
SiO_2	43·42	45·82
Fe_2O_3	17·64	15·20
Al_2O_3	19·00	25·00
Na_2O	3·66	4·12
K_2O	8·77	9·63
MgO	0·51	tr.
CaO	1·81	tr.
Mn	—	tr.
H_2O	4·33 ("loss")	0·41
	99·14	100·18

The difference in the amounts of Al_2O_3 is remarkable, but, as will be seen, the effect on the stopping power is not important.

A determination of the density afforded 2·963. Applying Bragg and Kleeman's method of computation (Bragg, 'Studies in Radioactivity,' 1912, Chap. V) it is found that 1 cm. in air corresponds to 0·000441 cm. in the mica according to Haughton's analysis, and to 0·000447 on Stone's analysis. It is evidently preferable to accept the latter value of the stopping power as being more certainly applicable to the mica in which the haloes occur.

A very considerable number of measurements were made on primary rings in the Mourne mica. Measurements on similar rings in the mica of Ballyellen were made under like conditions. Several observers took part in these measurements.

It was thought well to re-investigate the stopping power of the mica of Ballyellen. A new determination of the density gave a result a little higher than that appearing in the 'Philosophical Magazine' (April, 1910, p. 638). The density as recently found is 3·038. Two analyses are available: one by

Haughton cited by Dana ('Mineralogy,' p. 307, 5th edn.), and one, less detailed, by R. J. Moss, F.C.S. ('Phil. Mag.,' *loc. cit.*, p. 638). Applying the recent value for the density to Haughton's analysis, the stopping power comes out as 0·000450 cm., and on Moss' analysis 0·000451 cm. The former determination of density would raise these values about 3 per cent. As the errors in the determination of the density of mica are more likely to lower than to raise the result, the recent value of the density would seem to be the more reliable.

The results of comparative measurements on primary halo rings in the Mourne and in the Ballyellen micas will presently be given. It will then be seen that there appears to be a small but real and definite difference in the radial dimensions of these haloes. In consequence of this the micas of Arendal and Ytterby were searched for primary halo rings.

These micas are extremely ancient. They are found in the Archæan rocks of Norway and Sweden. Scarcity of undeveloped haloes was, therefore, to be anticipated. Search specially directed to their detection resulted in the finding of a few suitable for measurement, both in the Arendal and the Ytterby mica. The stopping power of these micas has, therefore, to be considered.

The Arendal mica was found to have a density of 2·9927. An analysis is cited in Dana's 'Mineralogy' (5th edn., Ap. III, p. 79). The density of the Ytterby mica was found to be 3·0656. A recent chemical examination of this mica by Ivar Nordenskjöld* shows that there are distinct chemical differences between the more highly altered and the less altered types of this mica. The mica in which the haloes have been investigated is considerably altered. Both the analyses are given below.

	Arendal.	Ytterby Fresh.	Ytterby Altered.
SiO_2	38·89	32·23	27·68
Al_2O_3	14·53	15·97	15·48
Fe_2O_3	4·58	7·87	14·25
FeO	7·85	28·06	19·39
MnO	1·06	-	—
MgO	20·28	2·75	6·38
K_2O	10·08	7·93	0 57
Na_2O	0·40	1·91	1·69
H_2O	0·94	2·80	7·51
F	1·49	· ·	0·14
CaO	—	0·23	3·48
Moisture	—	0·31	—
Y_2O_3	—	·	1·54
$(Nb, Ta)_2O_5$	—	—	1·36
	100·10	100·18	100·11

---

* Bull. Geol. Inst. of Upsala,' V, IX.

The chloritisation of the fresh Ytterby mica, attended with the introduction of Ytterbium and other rare elements, is indicated by the two analyses given above. For the Arendal mica, the mica equivalent of 1 cm. range in air is found to be 0·000447 cm., and for the fresh Ytterby mica it is 0·000456. For the chloritised Ytterby mica it is 0·000436.

The stopping power of the four micas principally considered in this paper are:—

Mourne mica	0·000447
Ballyellen mica	0·000450
Arendal mica	0·000447
Ytterby mica	0·000436 to 0·000456

These results are probably reliable within 2 or 3 per cent. The chief risk of experimental error is in the determination of density. There may be local variations of stopping-power in the altered Ytterby mica. There are local variations of colour-density and local cloudiness. Certain remarkable appearances in these ancient micas and affecting the haloes therein have already been referred to.

In making the measurements the method described in the 'Philosophical Transactions' (*loc. cit.*) was in nearly all cases adhered to. The nucleus of these undeveloped haloes is very often apparent. When this is the case the radius and the diameter of the halo were very often both read. Placing the halo a little ex-centrically in the field, by rotating the stage one of the fixed lines of the micrometer eyepiece is brought either into tangential exterior contact with one side of the halo or as bisecting the nucleus. The travelling wire is now brought up to the opposite exterior arc of the halo ring. It may be adjusted to a nicety. The halo may then be shifted in and out of its position so as to study the accuracy of fit. It is generally desirable to make several readings on the same halo.

In every case the aim in adjusting the micrometer lines to the halo should be to bring one side of the line into exterior tangential contact with the halo ring. The lighting and focus are all important. For the denser micas a powerful focus lamp is required. For less dense micas the softer beam from an opal-glass tungsten lamp is better. It will be readily understood that in these observations the uncertainty is mainly confined to unconscious variations of procedure when haloes are being compared, for the boundary, of course, is not a hard and definite line. On this account nothing is gained by excessive magnification. Each observer should use the magnification he finds most suitable.

Some of the results contained in the following table were obtained by observers who were not aware of the object of the measurements. In some other cases the observer was not informed which halo was expected to read a less radial dimension. Only one observer failed to find a difference between the Mourne and Ballyellen haloes. He considered that his readings should not be regarded as reliable, owing to the state of his eyesight. His readings are excluded at his own request. Dr. H. Poole obtained only a very small difference.

The first table contains results comparative of Mourne and Ballyellen haloes. The initials of the observers refer to Mr. L. B. Smyth, B.Sc.; Dr. J. H. J. Poole; Prof. Gatenby, Sc.D.; Dr. H. Poole, and myself.

Table I.

	Mourne.			Ballyellen.		
Observer.	No. of Haloes.	Total No. of Readings.	Mean Values of Radius.	No. of Haloes.	No. of Readings.	Mean Value of Radius.
J.	5	5	0·0137	6	6	0·0150
J.H.J.P.	1	1	0·0143	3	3	0·0155
J.	7	20	0·0140	8	17	0·0154
L.B.S.	3	9	0·0135	3	9	0·0152
L.B.S.	-	--	—	2	9	0·0146
J.	1	7	0·0135	1	5	0·0152
L.B.S.	1	5	0·0136	1	11	0·0147
J.	1	5	0·0129			
L.B.S.	2	14	0·0141	--	—	--
L.B.S.	1	7	0·0130	1	6	0·0138
L.B.S.	—	--	--	1	17	0·0141 Very faint.
L.B.S.	-	--	--	1	12	0·0142 Less faint.
L.B.S.				1	7	0·0141 Faintest.
L.B.S.	1	17	0·0142	--	-	
L.B.S.	1	6	0·0142		--	
L.B.S.	1	3	0·0136		--	
L.B.S.	1	18	0·0139		-	
J.	1	3	0·0127	1	10	0·0146
J.	1	9	0·0130		—	
G.	1	6	0·0136	1	4	0·0146
G.	1	10	0·0131	1	8	0·0141
H.P.	1	20	0·0139	—		—
H.P.	1	20	0·0142	1	20	0·0144
		175	0·0135 mean.		144	0·0146 mean.

The earlier readings (the readings are cited in chronological order) of Mr. Smyth, the readings of Dr. J. H. J. Poole, all those of Dr. H. Poole and of

Prof. Gatenby were made in ignorance of the source of the haloes under examination. If all my own readings are excluded the respective means are: for Mourne haloes, 0·0136 on 126 readings, and for Ballyellen haloes, 0·0144 on 106 readings. It is noteworthy that the published mean for Ballyellen first-ring haloes ('Phil. Trans.,' *loc. cit.*, p. 61) is 0·0145, neglecting the nuclear correction. This is in satisfactory agreement with the result given above.

An estimate of the probable error of the arithmetic means was made on the assumption that each reading constitutes an observation, and that all such observations are of equal credibility. Using the usual formula, the probable error of the mean for Mourne haloes is ±0·000085 and for the Ballyellen haloes ±0·000095.

The next table refers to haloes in the Archæan mica of Arendal.

Table II.—Arendal Haloes.

Observer.	No. of Haloes.	No. of Observations.	Mean Radius.	Remarks.
J.	1	10	0·0155	Results are on two different magnifications.
J.	1	8	0·0152	,, ,, ,,
J.	1	6	0·0157	Very faint halo.
L.B.S.	1	11	0·0153	Faint. Two magnifications used.
L.B.S.	1	16	0·0162	Very strong.
L.B.S.	1	10	0·0161	Clear, not so strong as last. A Ballyellen halo on six observations now read 0·0148.
J.	1	6	0·0164	(Small inner ring present.)
J.	1	2	0·0149	
J.	1	1	0·0154	
J.	1	8	0·0160	Very clear. (Small inner ring present.)
J.	1	4	0·0153	Faint but clear The last five results obtained on particular good lighting. A Mourne halo substituted now read 0·0133.
J.	1	3	0·0157	,, ,, ,,
J.	1	3	0·0157	Stronger
J.	1	1	0·0153	Like last
J.	1	1	0·0151	Faint
J.	1	5	0·0158	Difficult.
J.	1	2	0·0160	More definite.
J.	1	5	0·0160	Strong.
		102	0·0157 mean.	

The next table contains results on the black mica of Ytterby. Various specimens of the mica were used. Some from the Geological Museum of Trinity College; one from South Kensington Museum; some recently purchased.

A number of experiments were now made in which the procedure was varied in several ways by changing the objective and substituting Arendal for Mourne haloes and vice-versa. To cite the whole of them would occupy much space. The results given are in no way selected; those omitted are in full agreement

Table III.—Ytterby Haloes.

Observer.	No. of Haloes.	No. of Observations.	Mean Radius.	Remarks.
J.	1	1	0·0161	
J.	1	7	0·0156	
J.	5	9	0·0160	Quite clear.
J.	2	2	0·0157	
J.	1	4	0·0162	Delicate.
J.	1	9	0·0161	Strong.
J.	1	9	0·0160	,,
J.	1	10	0·0164	Barely visible.
J.	1	6	0·0159	Very faint. A similar Mourne halo now read 0·0134 on 16 observations.
		57	0·0160 mean.	

and were made at a later date. The objectives used were Leitz Nos. 3 and 4. Various modes of lighting were tried. The observer was myself.

(1) Using Objective 3, five Arendal haloes gave readings from 0·0151 to 0·0161; best halo read 0·0153; two Mourne haloes read from 0·0133 to 0·0138.

(2) Using Objective 4, one Arendal halo gave 0·0160; one Mourne halo gave 0·0133.

(3) Objective 4, a Mourne halo, mean of eleven readings, gave 0·0136. Objective 3, same halo, mean of three readings, gave 0·0138.

(4) Six Arendal haloes:

	(*a*)	on 5	observations	gave 0·0152		Objective 3.
	(*b*)	,, 5	,,	,, 0·0151		
	(*b*)	,, 5	,,	,, 0·0151		Objective 4.
	(*a*)	,, 3	,,	,, 0·0149		
	(*c*)	,, 5	,,	,, 0·0151, faint		Objective 3.
	(*d*)	,, 2	,,	,, 0·0157, fainter		
	(*e*)	,, 4	,,	,, 0·0157, fairly strong		
	(*f*)	,, 4	,,	,, 0·0161, very faint		
A Mourne halo		,, 4	,,	,, 0·0133, delicate		

(5) Objective 3, a Mourne halo, very sharp and strong, large nucleus, on twenty-three readings read 0·0141, on one reading read 0·0133. Objective 4, fifteen readings read 0·0140.

(6) More recently I concluded that Mourne haloes could be read from 0·0133

to 0·0141, and Arendal haloes from 0·0149 to 0·0161, with no definitely perceptible errors of adjustment.

The mean of the comparative results numbered (1) to (5) as above is: Mourne haloes, 0·0136; Arendal haloes, 0·0154.

Rapakivi Haloes.

The Rapakivi granites are stated by Sederholm to be of Jotnian age ('Bull. de la Comm. Géol. de Finlande,' Helsingfors, 1907, No. 23, p. 93). P. Eskola refers "Rapakivi" and related granites to pre-Jotnian times, and above the Archæan ('Bull. de la Comm. Géol. de Finlande,' No. 40, p. 10). The age of the Jotnian is probably Torridonian or Keweenawan. Eskola's estimate would refer the Rapakivi granites to Huronian or Upper Huronian times. Both the estimates agree in dating the intrusion of this granite from times considerably more ancient than the Devonian, and much younger than the Archæan.

The mica of the Rapakivi granite is limpid and seemingly unaltered, of a deep brown colour: Haloes are abundant in some flakes, but good first-ring haloes are very scarce. They appear to be derived from Uranium in nearly all cases. What appear to be first-ring Thorium haloes exist in it—but are excessively scarce. These scale 0·0115 q.p. This stage in the development of the Thorium halo has not hitherto been observed ('Phil. Trans.,' *loc. cit.*, p. 68). The specimens of Rapakivi used in my observations are labelled as coming from Viborg.

The following measurements have been made on several different haloes:—

(*a*) Very faint but clear. Nucleus very small. On 5 readings $r = 0{\cdot}0150$ to $r = 0{\cdot}0154$ (J).

(*b*) A stronger halo, on six readings $r = 0{\cdot}0151$ (L.B.S.); on eleven readings $r = 0{\cdot}0153$ (J).

(*c*) Faint but quite measurable. On ten readings $r = 0{\cdot}0153$ (J). Ballyellen rings now read on thirteen readings 0·0145 (J).

(*d*) Rather difficult. On ten readings $r = 0{\cdot}0153$ (J). The mean of the above is $r = 0{\cdot}0153$.

A series of readings comparative of Mourne, Arendal, and Rapakivi haloes were made (by myself) under very good conditions of lighting, such as afforded confidence that the reading lines were being effectively applied in definite contact with the exterior of the halo. The following results were obtained:—

Arendal. – Clear, light rings. A reading of the micrometer head was obtained which was not variable by more than one division either way without visible error of judgment.

This gave $r = 0·0153$.

Mourne.—(1) A clear light ring. Micrometer lines just touching. One and one half scale divisions introduced visible error.

This gave $r = 0·0133$.

(2) A coarser and more deeply coloured ring with larger nucleus was measured. Its edge was more fuzzy. This might have been read without definite error from 0·0137 to 0·0141.

(3) A ring of very clear development read apparently, decisively 0·0133.

Rapakivi.—Clear, delicate, first rings gave a mean reading of 0·0149 ± 0·0004.

It was determined that 0·0133 was decisively inapplicable, the lines being clearly trespassing on the halo and *within* its boundaries.

Arendal and Mourne.—A similar set of readings, using a different microscope and rather higher magnification, gave Arendal 0·0155, Mourne 0·0134 to 0·0137.

Rapakivi.—Some Rapakivi measurements were made, reading the radius directly; the nuclei being sharp and small. This is, in some respects, the best procedure, especially when some minute foreign object occurs at the outside boundary of the ring which may be used for setting one of the lines.

A very faint first-ring halo on eleven readings gave $r = 0·0151$. Another on twelve readings gave $r = 0·0150$.

Using a Reichert comparative eye-piece it was just possible to detect visually a difference in size of the Arendal and Mourne halo rings; but the observation could not, in itself, be regarded as a decisive demonstration. An endeavour was made by photographing these haloes on the same scale of enlargement and then printing side by side to detect the difference in dimensions. It was found that the definition of haloes as seen in the microscope is superior to what they exhibit in the photograph.

Collecting those results which have not appeared in Tables I, II and III, we have:—

Mourne.	Arendal.	Rapakivi.
0·0136	0·0156	0·0152
0·0133	0·0153	0·0154
0·0136	0·0160	0·0153
0·0138	0·0152	0·0153
0·0133	0·0151	0·0153
0·0141	0·0151	0·0151
0·0133	0·0149	0·0150
0·0140	0·0151	—
0·0133	0·0157	
0·0133	0·0157	—
0·0136	0·0161	—
—	0·0153	
	0·0155	—
	0·0158	—
0·0136 mean	0·0155 mean	0·0152 mean

The results previously given are:—

Mourne	0·0135
Ballyellen	0·0146
Arendal	0·0157
Ytterby	0·0160

It will be seen that these results increase with the antiquity of the rock. It is held by some geologists that the Ytterby rock is older than the Arendal. It is not safe at this stage to claim more for these results (laborious though they have been) than that they favour the view that these primary Uranium rings are of greater radius in the older rocks. In so far they suggest that further observations are desirable.

In the appreciation of the foregoing results the nature and size of the nucleus is important. The rays do not proceed from a central point. An approximate correction for this is to assume that they are emitted from the surface of a sphere of two-thirds the radius of the nucleus (' Phil. Trans.,' *loc. cit.*). If the stopping power of the nucleus was the same as that of the mica, then the correction for the nucleus would be a deduction from the measured radius of the halo of a length equal to two-thirds the nuclear radius. Now in the Ballyellen haloes the nuclei are in general zircons. This substance possesses a stopping power greater than that of the mica as 3 : 2. This reduces the correction for the nucleus to half the radius of the nucleus. In the case of the Mourne haloes, however, the nuclei are seldom, if ever, zircons. They are composed of some substance of lower refractive index probably allenite or apatite. Now the stopping power of both these substances is nearly the same as that of the Mourne mica. (In allenite the mica-equivalent of 1 cm in air is 0·000429, and in apatite it is 0·000499.) Hence the correction remains sensibly at two-thirds the nuclear radius.

But not only does the Mourne halo in this way owe more of its radial dimension to the nucleus than does the Ballyellen halo, the Mourne nuclei appear to be in general appreciably larger than those in the Ballyellen mica. The reason for this probably resides in the fact that, in a mica so recent as the Mourne only a relatively large nucleus will be able to develop the initial halo ring. In not one instance was a completely developed halo detected in this early Tertiary mica.

As it is very difficult to be precise in the measurement of nuclei, I have thought it better to neglect this correction on all alike. There is a further reason for this procedure, in the fact that the nuclei of the very old haloes are often so blackened by radio-active staining as to be practically unmeasurable.

The Potential Difference occurring in a Donnan Equilibrium and the Theory of Colloidal Behaviour.

By A. V. Hill, F.R.S.

(Received September 27, 1922.)

(From the Physiological Laboratory, Manchester.)

J. Loeb, in a recent and stimulating work (1), has given a convincing, if somewhat over-emphatic, study of the colloidal behaviour of proteins in solution, based largely upon the theory of the Membrane Equilibrium first suggested by Donnan (4). In one important particular, however, his argument is incorrect. Loeb observed, by certain means (2) devised by himself, the potential difference (P.D.) between a protein solution on one side of a semi-permeable membrane and a solution of acid, or of acid and salt, on the other side. He found this P.D. to vary as the concentration of hydrogen ions, or of salt, was varied, in the same manner as did a number of other factors (osmotic pressure, viscosity and swelling). He found also that this P.D. could be "calculated" from the observed difference of p-H (or of p-Cl)* in the two solutions, on the basis of the theory of the Donnan Equilibrium, and he concludes that the excellent agreement between calculated and observed is a strong argument in favour of his explanation of other colloidal phenomena by that theory. This conclusion is not correct: the equality found by Loeb of the observed P.D., to that calculated from the difference of *p*-H is a necessary consequence of *any* mechanism which does not offend the Second Law of Thermodynamics, and in itself offers no support to the theory that the Donnan Equilibrium underlies the colloidal behaviour of protein solutions. That theory may rest on other and stronger ground; since, however, Loeb appears, throughout his book (and especially in Chapters VIII and IX) and in other places (2), (3), to lay great emphasis on this agreement of the observed P.D. with that "calculated" from the observed *p*-H's it is necessary to point out that this agreement proves no more than that the system investigated was in equilibrium, and that the observations were accurately made.

---

* If the hydrogen- (or Cl-) ion concentration be 10^{-x}N, then p-H (or p-Cl) = x.

Consider the case of gelatine chloride, inside a semi-permeable collodion tube under a pressure equal to the osmotic pressure, in equilibrium with a solution of HCl outside. The argument applies equally to any type of solution and to ions of any valency. Admittedly the P.D. calculated from the observed hydrogen ion concentrations, viz. $RT/F \log H_a/H_b$, (where H_a is the observed hydrogen ion concentration in the gelatine-chloride solution, and H_b that in the solution of HCl), agrees with the P.D. observed directly with calomel electrodes connected to the two solutions by saturated KCl. If it did not, the Second Law of Thermodynamics would be broken. Provided the system be in equilibrium, and whatever be the hydrogen-ion concentrations (as measured electrometrically) or their cause, it is inevitable that the P.D. observed by Loeb's method should be equal to that "calculated" from the difference between the electrometric determinations of *p*-H in the two solutions.

It is obvious that if an isothermal system be in a genuine state of equilibrium it is impossible by any means whatever to obtain an electric current by connecting reversible electrodes, which themselves form part of the equilibrium system, to two different points in it. Otherwise work can be obtained at a constant temperature from a system in equilibrium. For example, if two hydrogen electrodes be placed, one in the gelatine-chloride solution and the other in the HCl of the system considered, then, after a genuine equilibrium has been set up, the difference of potential between them must be *nil*, whatever be the difference of hydrogen-ion concentration.* It is impossible, therefore, to measure a difference of hydrogen-ion concentration directly by such means. On the other hand, calomel electrodes with saturated KCl to make contact with the solutions are *not* in equilibrium with the system: a real equilibrium in such a case can occur only after a sufficient time to allow the KCl to diffuse, or to equalize itself as far as possible, throughout the system, in which case (as with hydrogen or chlorine electrodes) the final P.D. observed will necessarily be *nil.* The saturated-KCl electrode measures, or attempts to measure, part only of the total difference of potential found with electrodes in equilibrium with the system, viz. that part which is due to the contact of dissimilar solutions across a boundary.

---

* Since this was written the author has had his attention drawn to an exactly similar argument by Michaelis [(5) p. 156].

Consider now the system shown in fig. 1, in which gelatine chloride (*a*) and

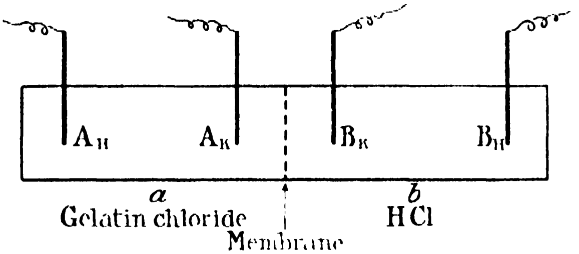


Fig. 1.

HCl (*b*) are in equilibrium of any kind across a semi-permeable membrane: in (*a*) is a hydrogen electrode, A_H, and a calomel electrode with saturated KCl, A_K; in (*b*) are two similar electrodes, B_H and B_K. From the Second Law of Thermodynamics, as pointed out above, it is inconceivable that there should be a difference of potential between A_H and B_H, which are reversible electrodes in equilibrium with their respective surroundings in a system in equilibrium; otherwise work could be obtained at constant temperature merely by the carriage of gaseous hydrogen from B_H to A_H. Hence we may write,

$$\text{Potential of } A_H - \text{potential of } B_H = 0,$$

from which, by alternate subtraction and addition of the potentials at A_K and B_K,

$$(\text{Potential of } A_H - \text{potential of } A_K) + (\text{potential of } A_K - \text{potential of } B_K) - (\text{potential of } B_H - \text{potential of } B_K) = 0. \quad \text{I}$$

Now, in using the hydrogen electrode, together with a calomel electrode with saturated KCl, the difference of potential as measured experimentally is assumed to give the logarithm of the hydrogen-ion concentration *plus* a constant C.

Thus the quantity inside the first bracket is equal to (RT/F log H_a + C), and the quantity inside the third bracket to (RT/F log H_b + C). Hence the first and third terms together come to (RT/F log H_a − RT/F log H_b), *i.e.* to RT/F log H_a/H_b, and equation (I) shows that the P.D. directly observed, viz. (potential of A_K − potential of B_K), is necessarily equal and opposite to this quantity. But RT/F log H_a/H_b is called by Loeb the "calculated P.D.," while (potential of A_K − potential of B_K) is his "observed P.D." It is obvious, therefore, from equation (I), that Loeb's "calculated P.D." with its sign changed must necessarily be equal always to the P.D. observed with the calomel electrodes with saturated KCl. Indeed, equation (I) would still hold if the calomel electrodes were replaced by any other electrodes, provided that these others were used also in conjunction with the hydrogen electrodes. *It will be noted that the proof of this equality has never mentioned or implied any theory of the mechanism by which the P.D. is set up.*

Now Loeb has made a great number of experiments, in which the difference of potential between two such solutions as we are considering was measured with calomel electrodes with saturated KCl and then compared with the difference of potential "calculated" from the difference of *p*-H as observed electrometrically for each solution separately. The agreement he regards as strong evidence for the existence of the Donnan Equilibrium, and for his explanation of colloidal behaviour on the basis of it. As a matter of fact, the argument given above shows that all such quantitative comparisons of the P.D. directly observed (with calomel electrodes employing saturated KCl), with the P.D. "calculated" from the difference of *p*-H (observed with hydrogen electrodes and calomel electrodes of the same type), are quite inconclusive, and serve merely to verify the fact that the system is in equilibrium. In his book (Chapter VIII, pp. 135, etc.) Loeb continually emphasizes the supposed argument that the "quantitative agreement between the observed P.D. and the P.D. calculated on the basis of the Donnan theory" (*in reality calculated solely by subtracting the P.D. observed with the hydrogen and calomel electrodes in one solution from the P.D. similarly observed in the other solution*) "leaves little doubt that the observed P.D. is exclusively determined by the Donnan Equilibrium"; and in a more recent paper he again emphasizes the same argument (3). In point of fact this type of comparison gives no verification whatever of any theory of the mechanism by which the observed P.D. is set up.

Imagine that for any physical reason whatever, not *necessarily* because of the existence of a Donnan Equilibrium involving an indiffusible ion, there is a difference of potential between two solutions in equilibrium with one another.* By a converse of the general thermodynamical argument given above it is then possible to prove that there must necessarily be a difference between the hydrogen- (or chlorine-) ion concentrations on the two sides of the membrane, and that this must be given by the equation RT/F log H_a/H_b. The fact that the difference of the *p*-H's can be calculated from the observed P.D. is no proof of any particular theory of the mechanism by which the P.D. is produced. Given that there *is* a P.D., the difference of *p*-H must necessarily result whatever be the explanation of the P.D. It is possible that another investigator may find, in some cases at any rate, that another theory of the occurrence of a P.D., say between the inside and the outside of a colloidal micella, is preferable to that of the Donnan Equilibrium: if so, it will still be necessary that the P.D. observed should be equal to that "calculated" in Loeb's way from the difference of the observed *p*-H's.

By a generalization of the argument given above it can be proved that in any two watery solutions of any kind, in equilibrium with one another, and containing *any* common ion A capable of passing from one to the other, the P.D. between the solutions is given by the quantity RT/F log A_1/A_2 where A_1 and A_2 are the concentrations of A in the two solutions respectively. This fact may be of value in the analysis of fluids, *e.g.* blood or serum, for ionic concentration. A collodion tube inserted into the fluid, and filled with water, would attain an equilibrium, and the concentration of the ions inside it could then be measured electrometrically or by analysis; while the observed P.D. would give the ratio A_1/A_2, from which, and from the observed value of A_2, A_1 could immediately be calculated.

Conclusion.

It is contended by Loeb that the Donnan Membrane Equilibrium, involving the presence of an indiffusible ion, is the basis of the colloidal properties of a protein solution. While the possibility of this conclusion is admitted, it is pointed out that one of the chief arguments employed in its favour by Loeb is incorrect. Loeb shows that the P.D. observed experimentally between a protein and a non-protein solution separated by a membrane agrees very exactly with that "calculated" from the difference in hydrogen-ion concentrations also observed experimentally, and concludes that this supports his theory. As a matter of fact, this equality is a necessary and inevitable consequence of the manner in which his observations were made and of general thermodynamical reasoning, and its proof is independent of any theory of the mechanism by which the P.D. is produced.

---

* Various causes of electrical potential differences in systems in equilibrium have been discussed by Michaelis (5) in the later chapters of his recent book.

REFERENCES.

(1) J. Loeb, 'Proteins and the Theory of Colloidal Behaviour,' McGraw-Hill Book Company, 1922.
(2) J. Loeb, 'Journal of Gen. Physiol.,' III, p. 667, 1920-21.
(3) J. Loeb, 'Journal of Gen. Physiol.,' IV, p. 769, 1922.
(4) Donnan, 'Zeitschr. f. Electrochemie,' 17, p. 572, 1911.
(5) Michaelis, 'Wasserstoffionenkonzentration,' Springer, Berlin, 1922.

The Periodic Opacity of Certain Colloids in Progressively Increasing Concentrations of Electrolytes.

By J. Holker, D.Sc., M.D.

Communicated by Prof. A. V. Hill, F.R.S. Received August 22, 1922.

(From the University, Edinburgh.)

I. Introduction.

In the course of an investigation of the complex interaction of an electrolyte, an emulsoid colloid and a suspensoid colloid (1921 (1), 1922 (2)), it was discovered that, as the concentration of the electrolyte present was progressively increased, the opacity of the liquid containing the three re-agents increased and decreased in a remarkable periodic manner. Considering the extreme complexity of the heterogeneous system in which the phenomenon was first observed, it was obvious at the outset that a very great deal of work would have to be done, and that the best method of approaching the problem would be by steadily simplifying the conditions of the experiments, and so by a process of elimination retain only those which were concerned in producing the phenomenon. Fortunately a suitable apparatus (referred to below) was available, with which it was possible to make a very wide range of experiments with a minimum amount of material.

II. The General Method of Investigation.

The method of measuring the opacity of the liquids studied was that described in an article recently published by the author (1921 (3)). It was early found that the purity of the water used for making solutions and dilutions was of great importance, since the reactions were so delicate that even the traces of salts found in tap water affected the results. A suitable and very efficient form of still was found to be that designed by Bourdillon (1913 (4)). The "conductivity" water from this was used as fresh as possible and stored, when necessary, in Jena glass flasks, with the exclusion of carbon dioxide and bacteria from the air. Again, to prevent, as far as possible, any reaction from taking place between the water and the glass vessels, only Jena glass test-tubes and flasks were used. The cleaning of the test-tubes was a source of trouble and error in earlier experiments, but later this was obviated by means of the following technique: Immediately after the completion of each series of experiments, the test-tubes were cleaned by washing, first with dilute acid or alkali, followed by several rinsings with "conductivity" water, and then steamed in an apparatus designed by the author. The particular advantages of this steaming apparatus lay in the facts that the steam used contained no traces of calcium, and that, after the preliminary washing with acid and water, there was a minimum of handling of the tubes. The test-tubes were placed in a series of racks, which were dropped into slots over the steamer. When sufficiently steamed the racks of tubes were removed to an electric drying oven and dropped into other slots. It may be mentioned here that it was found convenient to provide the Jena glass test-tubes with a lip, in order to facilitate pouring from the test-tube into the small opacimeter tube.

The colloid studied more particularly was human serum, which had been previously "inactivated" by heating to 56° C. for twenty minutes, a good supply of this type of emulsoid being readily available. Other colloids studied were various animal sera, gelatine, gum acacia, agar, gum mastic and dialysed hæmoglobin. The electrolyte used was principally sodium chloride, though potassium chloride, calcium chloride, hydrochloric acid and caustic soda were also tried. In all cases the electrolytes were certified by the makers as being pure for analysis.

The general procedure in carrying out each series of experiments, was to mix the reagents in the required proportions, and then to place the rack of test-tubes in a thermostat at 40° C. and shake the tubes initially and every hour afterwards. At the end of four hours, by which time equilibrium was attained, the opacity of the contents of the respective tubes was determined. The method of mixing the reagents in the required proportions may be exemplified by describing the routine employed in testing the effect of progressively increasing concentrations of sodium chloride on serum. A concentrated solution of sodium chloride is first prepared in "conductivity" water, and standardised against silver nitrate. From this is prepared a series of dilutions with conductivity water in large resistance glass test-tubes of about 80 c.c. capacity. The dilutions of sodium chloride having been prepared, 0·5 c.c. serum was then pipetted into each of a series of Jena glass test-tubes, and 2·0 c.c. of the required dilution of sodium chloride added. In this way the serum was diluted five times with a solution of sodium chloride which progressively increased in concentration in the successive test-tubes. But, in adding the sodium chloride to the serum, the saline itself was diluted. To correct this error in the example given, the concentration of the sodium chloride added to the serum was multiplied by the factor 0·8, which gave the net concentration of sodium chloride in the mixture, neglecting, of course, that originally present in the serum. An important factor influencing the degree of turbidity of the "solutions" was the amount of shaking given to the test-tubes. Some shaking was necessary in order to ensure the thorough mixing of the electrolyte and emulsoid. It was, however, found necessary to keep the amount of shaking to as near a minimum as possible, since frequent shaking markedly increased the degree of turbidity. Uniformity of shaking was facilitated by using a test-tube rack so constructed that by shaking the rack one shook the whole of the tubes at the same time and to same extent.

In order readily to study the individual periodic opacity curves, and also to compare the different curves one with another, it was found desirable to give the oscillations a numerical value, and, since the characteristics of an oscillation depend on its amplitude (A) and on its length (L), a convenient measure was found to be the ratio, A/L, between these two variables. Such a ratio was termed a characterising ratio. A complete oscillation was taken to be that part of a curve which lay between two successive minima. The amplitude of the oscillations was measured in degrees of opacity, and was taken as the length of the perpendicular dropped from the maximum to the line drawn through the two minima of the period. The length of the oscillation was measured by the difference in the normality of the net amount of added electrolyte at the two minima marking the limits of that oscillation. The advantage of expressing the length in normalities lay in the fact that the action of different electrolytes on one and the same substance could be compared. Moreover, under the conditions outlined, the characterising ratio of a period could be defined as the theoretical amplitude that period would have, if its length were represented by one equivalent amount of the added electrolyte.

III. The periodic opacity of human serum in progressively increasing concentrations of sodium chloride, and the influence of certain factors thereon.

In the present section it is proposed to describe in detail experiments dealing with mixtures of the electrolyte sodium chloride and the emulsoid human serum. The experiments were arranged in four series : (1) the periodic opacity curve of serum diluted with varying concentrations of sodium chloride, and the effect of varying (2) the wave-length of the incident light ; (3) the concentration of the serum ; and (4) the temperature. The serum used had been "inactivated" by heating for twenty minutes at 56° C.

1. *The periodic opacity curve for human serum diluted five times with concentrations of added sodium chloride ranging from* 0 *to* 31 *per cent.*– To cover the ground effectively in this series it was found necessary to do 118 experiments, and the work had therefore to be spread over several days. The diluted serum was allowed to stand in Jena glass test-tubes for four hours at 40° C., with shaking every hour. In fig. 1 the opacity of the five times diluted serum was plotted against the net concentration of sodium chloride after mixing with serum, neglecting, of course, that originally present in the serum. These curves show that the opacity of serum in the presence of increasing concentrations of sodium chloride varies in a markedly periodic manner. With low concentrations of saline from 0 to 0·96 per cent. the opacity of the serum became less and less until it reached a minimum. Beyond 0·96 per cent. saline it increased steadily to a maximum at 2·24 per cent. and so on. Allowing for the

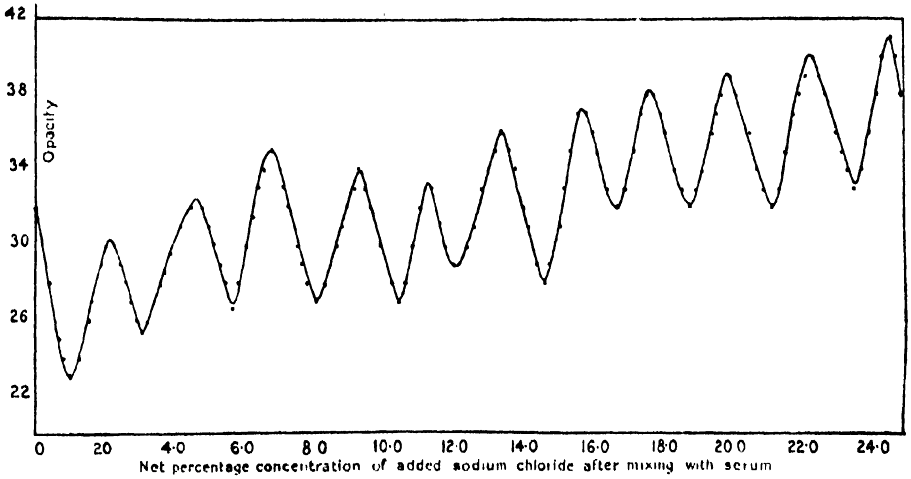


FIG. 1.—The opacity of human serum diluted five times with progressively increasing concentrations of sodium chloride.

fact that the experiments took several days to work out and that different batches of pooled serum were used on each day, the eleven oscillations produced were regular in their amplitude and frequency. Both the maxima and minima, however, increased in height with the concentration of the sodium chloride. The average characterising ratio of the whole curve, *i.e.* the average of the ratios of the amplitude to the length of the periods as defined in Part II, was found to be 16·5.

2. *The effect of varying the wave-length of the incident light.*—In the routine determination of the opacity of a liquid with the method described by the author, the incident light was of a pale yellow hue. In the present series of investigations it was sought to determine the effect of varying the colour and the wave-length of the incident light. For the purpose of testing merely changes of colour in the incident light, Wratten's light filters, adapted for use with the microscope, were found useful in that they provided a range of colour from dark blue to dark red. The screen was placed over the aperture in the shelf supporting the wire, and thus lay between the wire and the bottom of the small opacimeter tube. One and the same standard suspension was used for all the screens. This was the one usually prepared at the beginning of each day for the routine determination of a number of turbidities. It was therefore adjusted so as to have a reading of 200 without the interposition of any of the Wratten screens. The advantage of having one standard for all the screens lay in the fact that comparisons of the effect of changing the colour of the incident light were made under constant conditions. The routine standard having been prepared, its readings were retaken with the interposition of the respective screens, and the new readings plotted against the predominant wave-length of the light. The curve for these results is shown in fig. 2, and is seen to pass through a maximum where the incident light is yellow. It would appear then that in the determination of opacities by the method used, some other factor is involved besides the scattering of light. As the curve is very similar to that for the luminosity of the spectral colours, it is probable that with yellow light, to which the retina is most sensitive, one is able to distinguish the shadow space from the bright space with finer differences in intensity of the light emitted from these spaces than is possible with other colours. The readings of the standard for the respective screens having been obtained, the small opacimeter tube was filled with serum which had been diluted five times with progressively increasing concentrations of sodium chloride and allowed to stand for four hours at 40° C. with shaking every hour. The readings of the standard plus serum were then taken, and data thus

obtained which enabled one to calculate the depth the serum would have, when the shadow space was indistinguishable from the bright space. For example, when the colour of the incident light was dark blue, the standard emulsion alone had a reading of 120, and the standard plus serum diluted five times with distilled water had a reading of 77·35. The difference between the two readings was therefore 120—77·35 = 42·65. Therefore 42·65 mm. of standard were of equivalent turbidity to 50 mm. of the diluted serum. The length of the small opacimeter tube was 50 mm. Therefore, proportionally, 120 mm. of standard—the end point of the standard with the screen interposed —would be equal to $\frac{120 \times 50}{42 \cdot 65} = 140 \cdot 7$ mm. of serum. The figures thus obtained

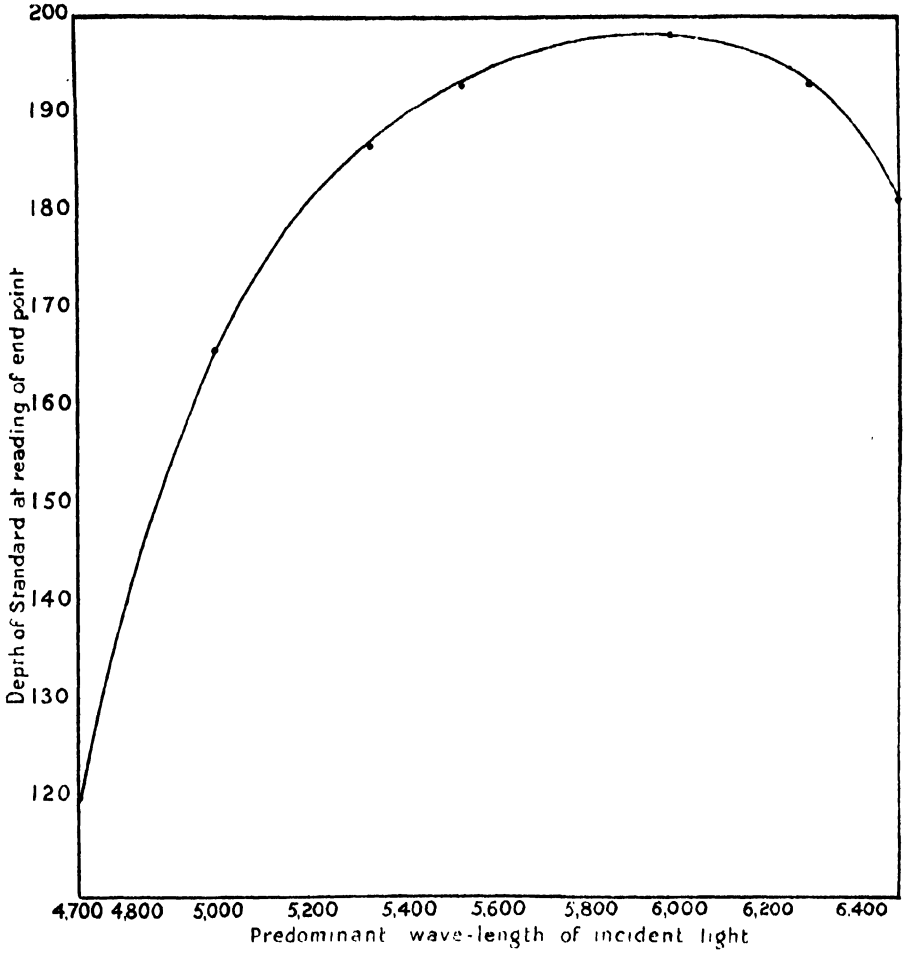


FIG. 2.—The influence of varying the colour of the incident light on the readings at the end-point depth of the standard emulsion.

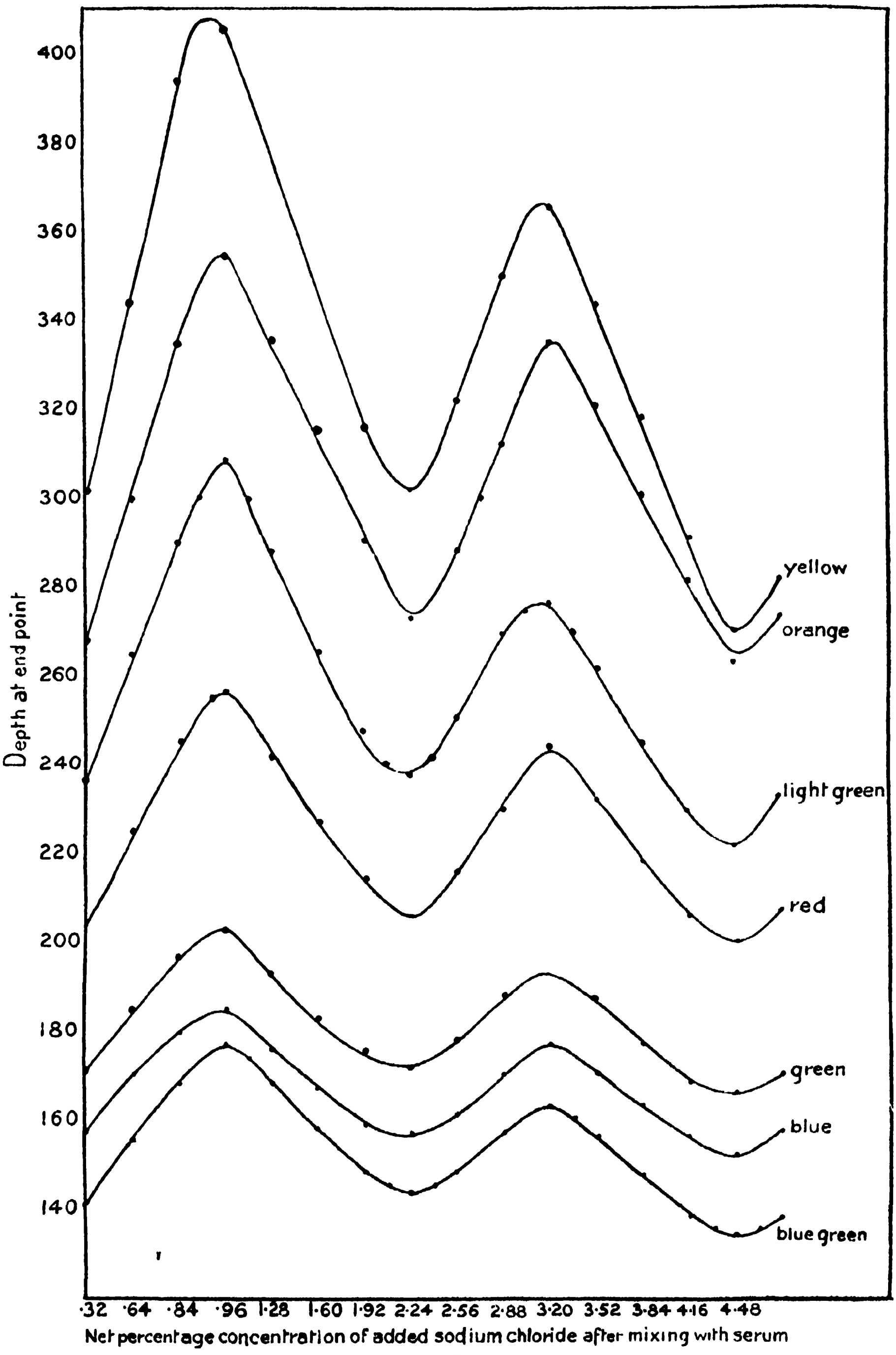


FIG. 3.—The effect of varying the colour of the incident light on the readings at the end-point depth of serum diluted five times with progressively increasing concentrations of sodium chloride.

for each series were then plotted against the concentration of sodium chloride added to the serum. The curves are shown in fig. 3. It will be seen that while the general level of the curves and the amplitude of the periods varied with the colour of the incident light, the length of the corresponding periods was remarkably constant. These experiments, therefore, suggest that the periodic opacity of serum in progressively increasing concentrations of sodium chloride cannot be due to interference phenomena, since in the latter case the length of the periods would have varied with the colour of the incident light. In order to follow better the influence of change of colour in the incident light on the reading, two sets of points were taken from the curves in fig. 3. One set represented the end-point depths of serum diluted with distilled water, and the other those for serum diluted with 1·2 per cent. NaCl. These two sets of points were plotted against the predominant wave-length of the incident

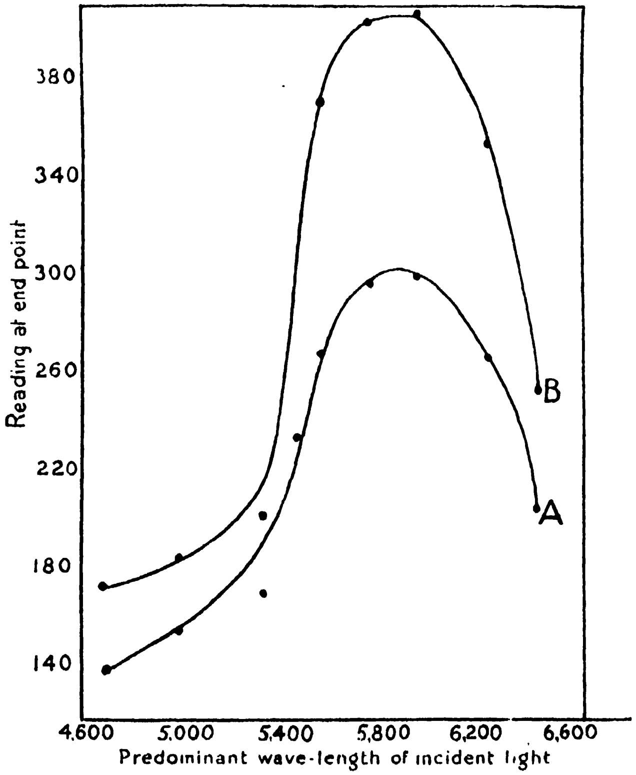


FIG. 4.—The effect of varying the colour of the incident light on the opacity of serum diluted (A) with diluted water, (B) with 1·2 per cent. sodium chloride.

light. The curves are shown in fig. 4, and are seen to be similar to that for the luminosity of the spectral colours. A practical point brought out by this series of experiments is, that the diffused pale yellow colour used in the opacimeter for contrast with the shadow is the most suitable colour for the purpose.

The screens used to obtain the foregoing results, while exceedingly convenient in enabling one to cover the whole range of colours in the visible spectrum in one and the same series of experiments, did not, however, render the incident light monochromatic, and, as it was very desirable to test the effect of varying the wave-length of the light, the experiments were carried further. For this purpose it was found necessary to modify the technique. In the standard apparatus the source of light is that of a metal-filament lamp, which is placed directly under the opacimeter tube. Between the lamp and the lower end of the opacimeter tube is placed a sheet of opal glass which serves to scatter the light from the incandescent filaments, and to provide a source of diffused pale yellow light of nearly uniform intensity. In the present experiments the lamp and sheet of opal glass were discarded and in their place was fixed a mirror arranged at an angle of approximately 45 degrees to the vertical. In front of the mirror was the source of light. Between the source of light and the mirror was a lens which was so adjusted that a horizontal beam of light was made to fall on the mirror, which then reflected the light vertically up the opacimeter tube. The monochromatic lights used were those of the red from the lithium flame (6708 A.U.), the yellow from the sodium flame (5893 A.U.), and the green from the mercury vapour lamp (5461 A.U.). Approximately monochromatic light was also used with the blue violet (4358 A.U.) and deep violet (4078 and 4046 A.U.) from the mercury lamp. In each case the light reflected by the mirror up the opacimeter tube was tested before use with the aid of a spectroscope. The incident light of wave-length 5893 A.U. was obtained by fusing thin-walled glass capillaries in the Bunsen flame. The light of wave-length 6708 A.U. was obtained by vapourising a saturated solution of lithium chloride by means of a sparking coil, one terminal of which being below and the other just above the surface of the solution, and then blowing the vapour by means of a gentle current of air into a Bunsen flame. The light from this flame was not of itself quite monochromatic, but was readily rendered so by placing between the flame and the mirror a sheet of the ruby glass used in photography. The pure monochromatic light of wave-length 5461 A.U. was obtained by placing a Wratten filter 77A between a mercury vapour lamp and the mirror. The approximately monochromatic light of wave-length 4358, 4078 and 4046 A.U. was obtained by placing a dark blue screen between a mercury vapour lamp and the mirror. The screen was made by impregnating a photographic plate with a solution of a dyestuff. On examining with the spectroscope the blue light reflected by the mirror, it was found to contain a well-marked band in the blue violet (4358 A.U.), two faint bands in the deep violet (4078 and 4046 A.U.), a very indistinct band in the green, and another very indistinct band in the yellow. In all cases the experiments were carried out in a dark-room, and, as an additional precaution, the apparatus was so screened that no light other than that from the monochromatic source could fall on the mirror. Moreover, to eliminate any personal factor, the determinations of the end-point were made by one observer, and the readings of the scale of the apparatus measuring that end-point were made by another observer. The results obtained in these experiments are summarised in Table I.

Table I.

The positions with respect to the abscissa (the net percentage concentration of added NaCl after mixing with serum) of the first three minima and maxima of the opacity curve of "inactivated" human serum diluted five times with progressively increasing concentrations of NaCl. Temperature, 40° C. Time, 1 hours, with shaking every hour.

Wave-length of incident-light.	Minima.			Maxima.		
	1st.	2nd.	3rd.	1st.	2nd.	3rd.
6708 A.U. (red) … (1)	1·00	3·20	5·60	2·20	4·60	6·80
(2)	0·95	3·25	5·65	2·25	4·45	6·85
5893 A.U. (yellow)	1·00	3·20	5·55	2·15	4·55	6·85
5461 A.U. (green)	0·95	3·20	5·55	2·15	4·55	6·85
4358 A.U. (blue violet) 4078 A.U. }(deep violet) 4046 A.U. } (1)	1·00	3·20	5·60	2·25	4·55	6·80
(2)	0·95	3·15	5·60	2·25	4·50	6·80
Average	0·98	3·20	5·59	2·21	4·55	6·83
Mixed light from metal filament lamp	0·96	3·20	5·60	2·24	4·60	6·75

It will be seen from Table I that the positions of the minima and maxima with respect to the abscissa (the net concentration of added sodium chloride after mixing with serum) remain constant for varying wave-lengths of incident light. It follows from this that the length of any given oscillation remains constant for varying wave-lengths of incident light, and that, therefore, the periodicity in the opacity is not due to an optical interference of the light scattered by particles uniformly changing in size, but is rather due to an oscillatory change in the physical condition of the particles themselves.

3. *The effect of varying the concentration of the serum.*—In this series of experiments it was sought to determine the effect of varying not only the concentration of the saline but also that of the serum as well, and, as this meant the covering of a very wide field, the upper concentration of the sodium chloride added was limited to 10 per cent. The conditions of the experiments were the same as those for subsection 1. The dilutions of serum varied from 1 in 1·5 to 1 in 800. The results were plotted in fig. 5. It will be seen that

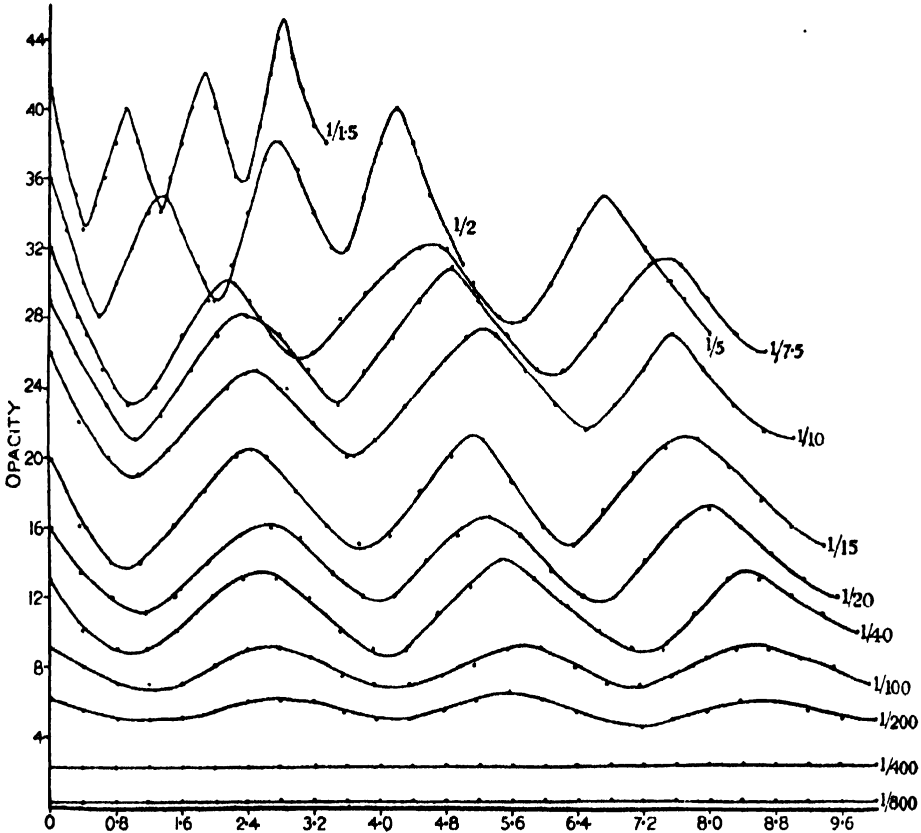


FIG. 5.—The opacity of human serum diluted with progressively increasing concentrations of sodium chloride.

the periodic character of the curve persisted to a high degree of dilution of the serum, and that both the amplitude of the oscillations and their number increase with the concentration of the serum. To bring this point out more closely the characterising ratio was determined for each of the curves in fig. 5. The characterising ratios were then plotted against the concentration of the serum (see fig. 6).

It is evident that the characterising ratio of the curve for serum diluted not more than forty times (2·5 per cent. serum or 0·212 per cent. serum protein) is

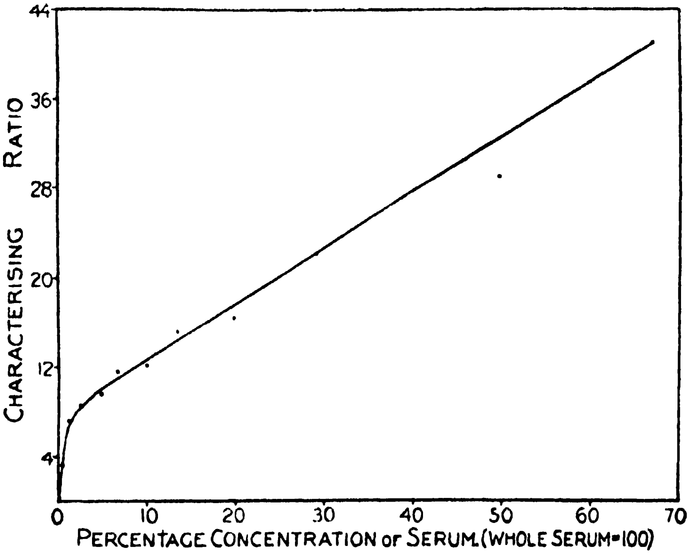


FIG. 6.—The characterising ratios for human serum diluted with sodium chloride.

a simple linear function of the concentration of the serum. With serum diluted more than forty times, the ratio diminishes very rapidly.

4. *The effect of varying the temperature.*—As it was well known that protein solutions, when heated, aggregate, and at higher temperatures precipitate, it was sought to obtain some knowledge as to the range of temperatures for serum within which the periodic phenomenon was observed. The serum was diluted five times with progressively increasing concentrations of sodium chloride up to 10 per cent. added, and allowed to stand at the required temperature for four hours, with shaking every hour. The curves for the temperatures 5°, 20°, 30°, 40° and 56° C. are shown in fig. 7. In the ice chest at 5° C. the influence of the added saline up to 10 per cent. on the diluted serum was almost negligible and no periodicity of opacity observed. At the temperatures 20°, 30°, 40° and 56° C. the periodic phenomenon was well marked, and the amplitude of the periods increased with the temperature, so that at 56° C. it was very pronounced. On the other hand the length of the corresponding periods remained the same for all the temperatures. In some further experiments, it was found that at 65° C. the amplitude of the periods was still more marked, but that there was a marked falling off in the regularity with which the points fell on the curve. At 75° C. the oscillations were marked

but irregular, and when the experiments were repeated the curve for the second series was unlike that for the first. These two curves for 75° C. had one point in common, however, in that apart from the oscillations they both showed a general rapid increase of opacity with increasing concentrations of added saline, so that at higher concentrations of added saline the turbidity of serum was very marked and precipitation set in. In order to arrive at a clearer understanding of the effect of temperature, the characterising ratios were determined for the curves of fig. 7, and then plotted against the respective temperatures. The curve thus obtained is shown in fig. 8. It will be seen that the rate at

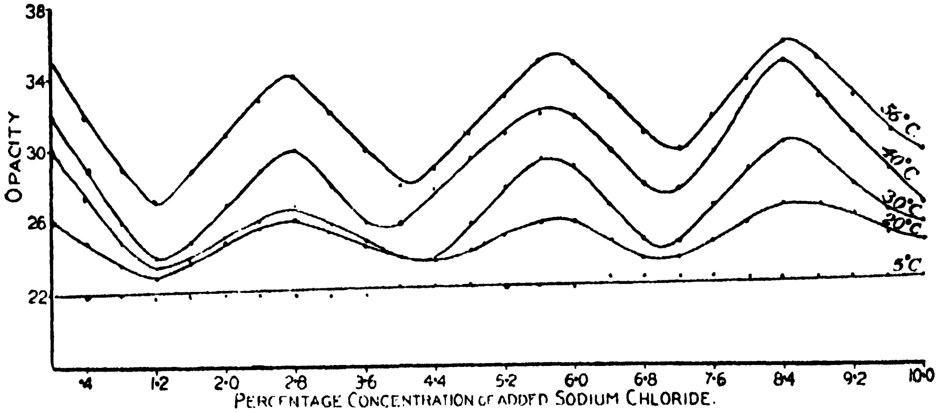


Fig. 7.—The effect of temperature on the opacity of human serum diluted with progressively increasing concentrations of sodium chloride.

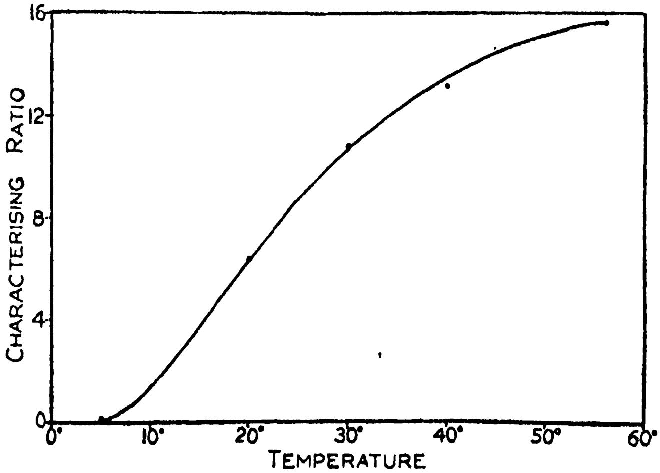
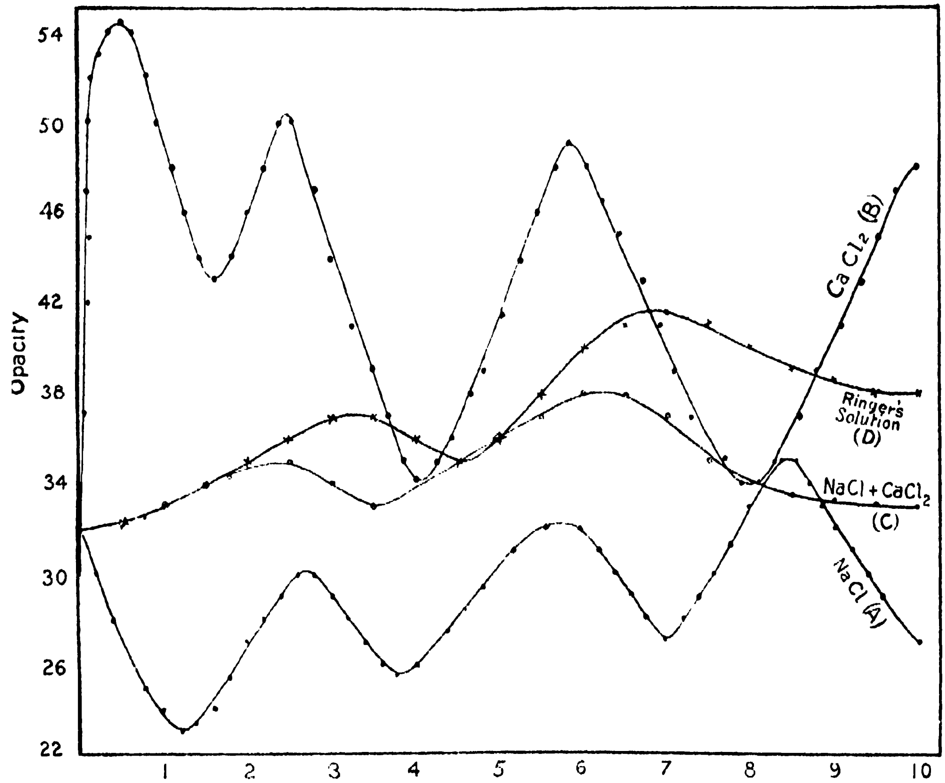


Fig. 8.—The characterising ratio of serum diluted five times with progressively increasing concentrations of sodium chloride plotted against the temperature.

which the characterising ratio increases with respect to the temperature, increases more and more until it attains a maximum value at a point which corresponds to a temperature of 17·5° C., and then diminishes more and more, tending finally to become constant. In other words, from the temperature of the ice chest to room temperature, the amplitude increases more and more rapidly; from room temperature to 56° C. the amplitude increases more and more slowly; above 56° C. the amplitude would be nearly constant, provided no other changes took place in the serum. But we have seen that increasing the temperature increases the degree of aggregation of the protein particles, and, if carried out sufficiently far, ultimately leads to precipitation of the protein. It would appear, therefore, that the decreasing rate of increase of amplitude above room temperature is due to the increasing degree of aggregation of the protein particles, and that when the degree of aggregation has passed a certain limit the amplitude and length of the periods become irregular.

IV. *The opacity of serum diluted with various electrolytes.*

The author has shown in the last section that, when serum is diluted with progressively increasing concentrations of saline, its opacity exhibits a remarkable series of oscillations. It was advisable therefore to extend the investigation, and to test the effect of other neutral salts. Further, to form a basis of comparison for the action of the salts, it was also desirable to determine the effect of acid and alkali. The neutral salts taken were sodium chloride, potassium chloride, and calcium chloride, these being salts which are found in physiological fluids, such as serum itself. In addition two mixtures of these salts were tested, one containing sodium chloride and calcium chloride in the proportion of 0·85 NaCl to 0·022 $CaCl_2$, and the other containing NaCl, $CaCl_2$, and KCl in the proportions of 0·85 : 0·022 : 0·056. The latter mixture therefore was equivalent to Ringer's solution. The acid used was hydrochloric, and the alkali caustic soda. The general method of investigation was to dilute 0·5 c.c. of serum, which had been "inactivated" by heating at 56° C. for twenty minutes, with 2·0 c.c. of the solution of salt, acid or alkali respectively made up in progressively increasing concentrations, and then to place the test-tubes at 40° C. for four hours, with shaking every hour. The resulting turbidities of the serum were plotted against the concentration of electrolytes added and the curves corresponding to these results will be found in figs. 9 and 10. In fig. 9 the concentration of the mixed salts was expressed in terms of the contained sodium chloride in order to make it possible to compare the curves for serum plus mixed salts with that for serum to which sodium chloride alone was added. In fig. 10 the concentrations were expressed in normalities, in order to form a basis of comparison between the salts, the acid and the alkali.



A. Percentage concentration of added sodium chloride.
B ,, ,, ,, calcium chloride.
C. ,, ,, ,, sodium chloride in the mixture of NaCl and $CaCl_2$.
D. ,, ,, ,, sodium chloride in the mixture of NaCl, KCl, and $CaCl_2$.

FIG. 9.—The opacity of serum diluted five times with progressively increasing concentrations of sodium chloride, calcium chloride, and balanced solutions of the same.

On referring to fig. 9 it will be seen that the effect of sodium chloride and calcium chloride on the opacity of serum offered a striking contrast. With sodium chloride the initial change was a fall in opacity, with calcium chloride the initial change was a sharp increase in opacity. The mixed salt solutions, on the other hand, only feebly increased the opacity in the initial stage. The antagonising effect of calcium chloride on sodium chloride was therefore displayed by relatively small proportions of calcium. In their general upward tendency, however, the curves for the mixed salt solutions still resemble that for sodium chloride. The antagonising effects of sodium and calcium may be due in part to the action of their ions on the serum protein, since a similar phenomenon is found to occur with caseinogen. That the serum protein is probably not wholly responsible for the effect will be seen on considering the kinds of anions found in serum. These consist not only of chlorides, but also bicarbonates and phosphates. Sodium ions would therefore tend to keep the anions of serum in solution, whereas calcium ions would tend to form insoluble salts with both bicarbonate and phosphates. It will be seen from fig. 10 that the neutral salts examined all produced a periodic opacity of the serum with progressively increasing concentrations of salt. The curves for sodium and potassium chloride resembled each other in amplitude and frequency, and also in their general upward tendency. The characterising ratio for sodium chloride was 16·5 and for potassium chloride 20·0. It is of interest to observe

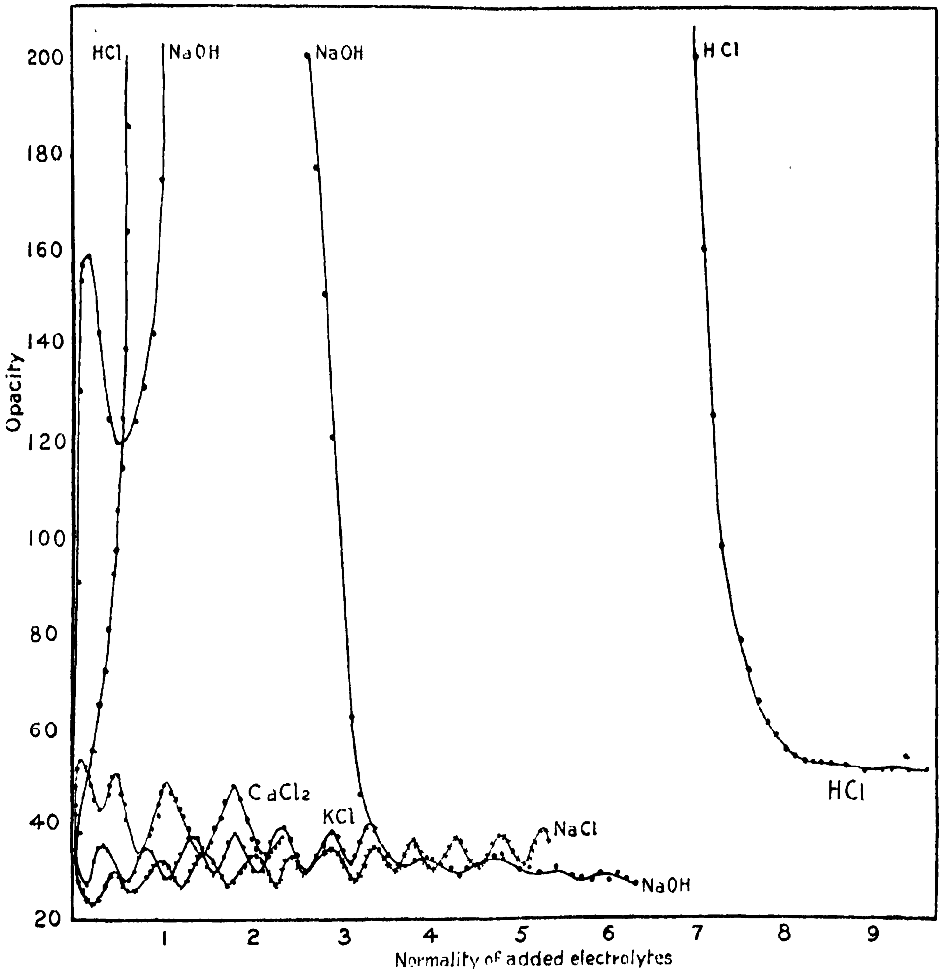


FIG. 10.—The opacity of serum diluted five times with progressively increasing concentrations of various electrolytes.

that the ratio 20·0 : 16·5 = 1·21 for potassium chloride and sodium chloride respectively, is very close to the ratio 131·2 : 110·3 = 1·19 for their equivalent conductivities at infinite dilution at 18° C. In the case of calcium chloride both the amplitude and the length of the oscillations were greater than with the other neutral salts. The line drawn through the maxima points for calcium chloride had a general downward tendency, in contrast to the similar lines for sodium and potassium chlorides which had an upward tendency. The characterising ratios were successively 15·0, 13·8, 12·2 and 10·4. The average characterising ratio 12·9 was therefore less than that for sodium chloride. Moreover the oscillations were markedly damped in character, whereas those for sodium chloride were not damped.

When the curves for the effect of hydrochloric acid and caustic soda are compared with those for the effect of sodium, potassium, and calcium chloride, it will be seen that the action of the acid and alkali was extremely great compared with that of the neutral salts. With the acid the oscillatory character of the curve was completely lost, and instead there was one maximum representing an extraordinarily high degree of turbidity compared with that for neutral salts, and also much higher than that for caustic soda. The macroscopic appearance of the serum in the test-tubes agreed well with those differences as given by the opacimeter. With hydrochloric acid in concentrations between 1 × N and 7 × N a dense white turbidity was produced. With neutral salts the differences in turbidity were often so slight as to be undetected when observed in the test-tube. With caustic soda the oscillatory character of the curve was not completely lost, since there was a fall in turbidity with concentration of NaOH between 0·2 N and 0·5 N.

A point brought out by these curves is that the effect of an electrolyte on serum appears to be highly specific.

V. *The periodic opacity of various colloids.*

A considerable number of detailed experiments were made with various animal sera, with gelatine, with hæmoglobin and with the carbohydrate colloids, gum acacia, agar, and gum mastic. A brief outline of the main results obtained from these experiments is given below.

It was found that all the sera examined exhibited the periodicity with progressively increasing concentrations of sodium chloride, the average characterising ratio for each curve being as follows: horse 4·35, rabbit 6·51, sheep 14·38, ox 14·9, man 16·5. It was of interest to observe that human serum had the highest characterising ratio.

It was found that gelatine, gum acacia, and agar all exhibited the periodicity with sodium chloride, and that the curves were characteristic for the respective substances. Of considerable interest was the fact that dialysed horse serum, dialysed hæmoglobin from the sheep and dialysed gelatine exhibited the periodicity with sodium chloride. In these cases, therefore, the reactions must have been in the main between the protein and the added electrolyte. Gelatine also reacted periodically to calcium chloride, hydrochloric acid and caustic soda, but the curves obtained were markedly different from those for serum with the same electrolytes. When, however, gelatine was dissolved in a solution of inorganic salts made up in the proportions found in serum, the periodic opacity curves with both sodium chloride and calcium chloride were similar to those for serum. An important series of experiments was one which showed that a very slight change in the initial hydrogen-ion concentration of the gelatine, whether on the acid or the alkaline side, had a pronounced effect on the constants of the periodic opacity curve for sodium chloride. It was also found that the value of the characterising ratio for different concentrations of serum, of gelatine or of gum acacia, as the case might be, was, over a wide range, a simple linear function of the concentration of the colloid. In this respect serum was six times more effective than gelatine, and twelve times more effective than gum acacia in increasing the value of the characterising ratio with increasing concentrations of colloid.

In Part III it was shown that with human serum the length of any given oscillation remained constant for varying wave-lengths of the incident light. As serum is a typical emulsoid colloid, it seemed advisable to extend the experiments with monochromatic light to a colloid which is regarded as a typical suspensoid. For this purpose gum mastic was chosen. It was found that, when the original solution of gum mastic in alcohol was diluted one thousand times with progressively increasing concentrations of sodium chloride, the degree of aggregation of the particles of the resulting suspension was so great that it was impossible to obtain a measurable turbidity. In order to prevent excessive aggregation of the particles forming the suspension, the surface tension of the system was reduced by means of alcohol. The original solution of gum mastic in alcohol was diluted one hundred times with alcohol. The new solution was then diluted ten times with progressively increasing concentrations of sodium chloride. The final dilutions of the gum mastic was therefore one in one thousand as before, but the suspension contained 10 per cent. of alcohol. In this way there was prepared a suspension in which the periodicity was not only measurable with the opacimeter, but was also visible to the naked eye.

It was found that the curves on the three consecutive days after preparing the one in one hundred dilution of gum mastic in alcohol were successively lower, but that after about the third day a position of equilibrium was attained. These experiments showed clearly that the dilution with alcohol of what was probably a saturated solution of gum mastic in alcohol, leads not only to a reduction in the concentration but also to a change in the physical condition of the " dissolved " gum mastic. During the time this change was taking place no alteration was visible in the one in one hundred solution, which appeared rather to maintain a constant high degree of transparency. For the purpose therefore of testing the effect of varying the wave-length of the incident light, a one in one hundred dilution in alcohol was prepared from the original solution of gum mastic in alcohol, and the new solution then allowed to stand at room temperature for about two weeks before suspensions were prepared. The methods of obtaining the monochromatic light were those described in Part III. The summarised results of these experiments are shown in Table II. It is of interest to mention that between the first and second series of readings for metal filament lamps there was an interval of thirty-eight days, the closeness of the results indicating the high degree of reproducibility of the phenemenon.

Table II.

The positions with respect to the abscissa (the percentage concentration of added sodium chloride) of the first two maxima and minima of the opacity curve of a gum mastic solution diluted one thousand times with progressively increasing concentrations of NaCl. Temperature 40° C. Time, four hours, with shaking every hour.

Wave-length of incident light.	Maxima.		Minima.	
	1st.	2nd.	1st.	2nd.
6708 A.U. (red)	1·5	6·95	4·60	8·10
5893 A.U. (yellow)	1·6	7·00	4·50	8·15
5461 A.U. (green)	1·6	6·80	4·45	7·90
4358 A.U. (blue violet) **4078 A.U.** } **(deep violet)** **4046 A.U.** }	1·6	7·00	4·60	8·10
Average	1·58	6·94	4·54	8·06
Metal light from metal filament lamp (1)	1·55	6·95	4·5	8·10
Metal light from metal filament lamp (2)	1·60	6·90	4·6	8·10
Average	1·58	6·93	4·55	8·10

The results shown in Table II show that with the suspensoid colloid, the positions of the maxima and minima are constant for varying wave-length of the incident light, and it is confirmed once more, therefore, that the periodicity cannot be due to optical interference of scattered light, and that it must be due rather to a periodic change in the physical condition of the particles themselves.

VI. The periodic opacity of certain mixtures of aqueous solutions of the inorganic salts found in serum.

The particular points of interest in the experiments outlined below were that none of the re-agents mixed together initially had any trace of a colloid in them, and that the effect of the presence of an organic substance, animal or vegetable, was completely eliminated.

The solution of serum salts used contained the following re-agents: NaCl 0·59 per cent., KCl 0·04 per cent., $CaCl_2$ 0·04 per cent., $MgCl_2$ 0·025 per cent., NaH_2PO_4 0·0126 per cent., and $NaHCO_3$ 0·351 per cent. Such a solution when properly prepared was clear, but in the presence of progressively increasing concentrations of calcium chloride exhibited the periodic opacity phenomenon. It was observed that the presence of the potassium chloride and sodium chloride in the solution of the serum salts increased both the solubility of the reaction products and characterising ratio of the curve. The presence of magnesium chloride, on the other hand, diminished both the solubility of the reaction products and the characterising ratio of the curves. On successively eliminating the re-agents making up the solution of serum salts, it was finally found that the periodic opacity phenomenon was exhibited by a mixture of sodium dihydrogen phosphate and sodium bicarbonate in the presence of progressively increasing concentrations of calcium chloride. It was not exhibited, however, by either the bicarbonate or acid phosphate alone, nor when disodium hydrogen phosphate was similarly treated. Continuing the process of simplification it was finally found that the phenomenon was exhibited when a "buffer" solution of acid and alkaline sodium phosphate was treated with progressively increasing concentrations of calcium chloride (see fig. 11).

VII. The theory of the periodic opacity phenomenon.

Periodic opacity might be due to optical interference of the light scattered by particles which uniformly changed in size with progressively increasing concentrations of added electrolyte. If so, then experiments made with

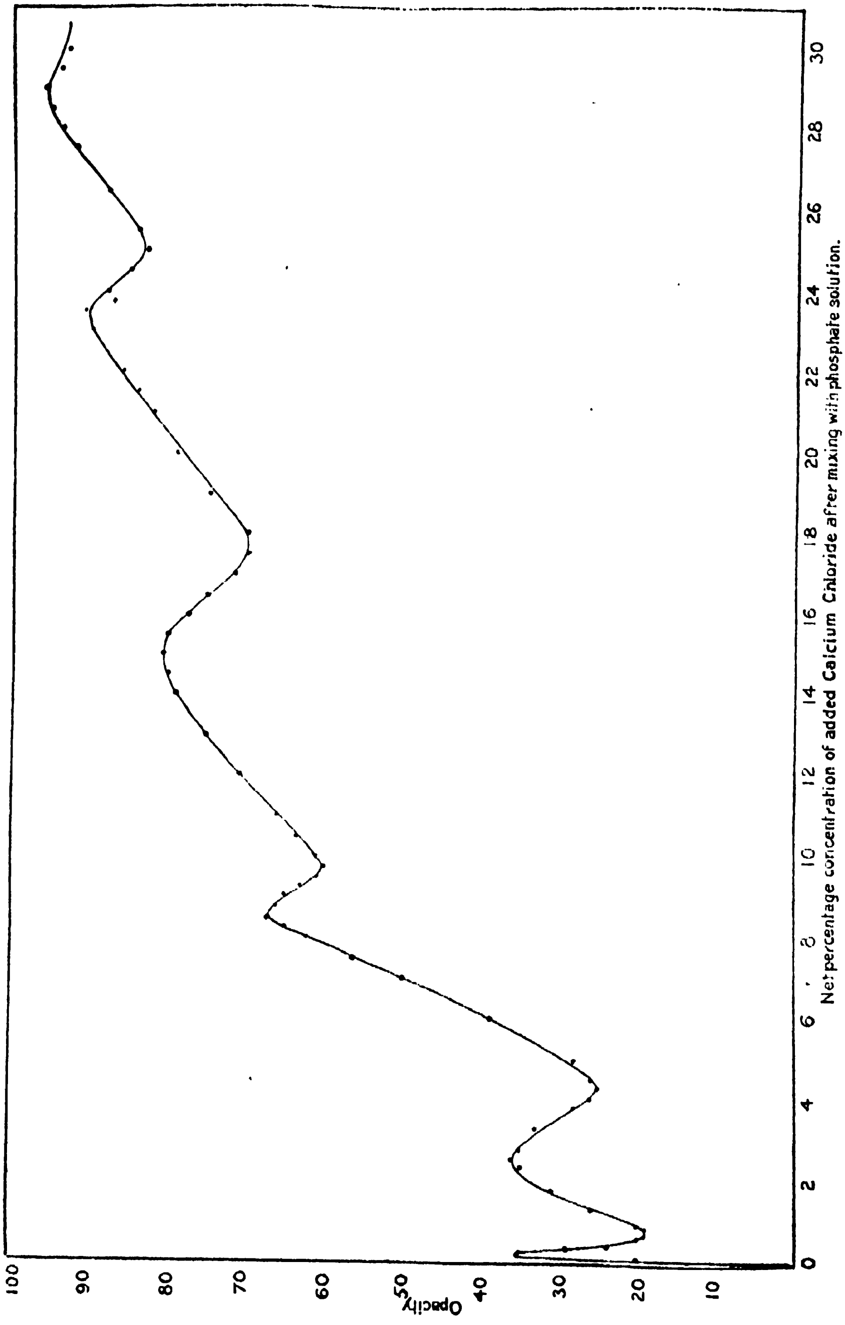


Fig. 11.—The periodic opacity of a phosphate "buffer" solution in progressively increasing concentrations of calcium chloride.

varying wave-lengths of incident light would produce curves in which the length of a given period would vary as the wave-length of the incident light. But it was shown in Sections III and IV with typical emulsoid and suspensoid colloids that the length of a given period remains constant for varying wave-lengths of incident light. Therefore, the periodicity in the opacity is not due to optical interference of scattered light.

If the phenomenon, on the other hand, is due to a periodic variation in the properties and physical condition of the particles scattering light, it should be possible to devise experiments in which the phenomenon is visible to the naked eye without the aid of a sensitive apparatus. But it was shown (1922 (2)) that a heart extract could be made to exhibit a periodic flocculation of its particles in the presence of an electrolyte. Moreover, the periodic flocculation became more and more marked with increasing concentrations of suspensoid. Again, it became less and less marked with diminishing concentrations of suspensoid, until finally it was not visible to the naked eye. On continuing the dilutions, the periodicity could still be detected with the aid of the opacimeter, and the positions of the maxima and minima still agreed with those obtained with more concentrated suspensoids. Significant also was the fact that in the region between the extremes of highly concentrated and highly diluted suspensoid, where the periodic opacity phenomenon was both visible to the naked eye and measurable with the opacimeter, the positions of the maxima and minima for visible flocculation agreed with those obtained with the opacimeter. In Section V was shown a second type of reaction in which a visible periodicity of opacity was obtained with the typical suspensoid colloid, gum mastic, in which the tendency to aggregation of the particles forming the suspension was so great that the oscillations could only be measured with the opacimeter when the surface tension of the system was reduced by the presence of a sufficient amount of alcohol. Even without alcohol, however, there were visible oscillations in the degree of aggregation of the particles. Again, the periodicity in opacity was visible to the naked eye in some experiments with water-glass and hydrochloric acid in which the physical condition of the reaction product was a gel. The foregoing examples provide direct evidence that the phenomenon is due to a periodic variation in the physical condition of the particles of the colloid scattering the incident light, and it follows, therefore, that the phenomenon is due to a periodic reaction between the colloid and the added electrolyte. To explain this periodic reaction it was first sought to apply those physico-chemical properties of matter, such as hydration of the colloid and the ions, and their electrolytic and hydrolytic dissociation. It was found, however, that none of these properties seemed capable of being adapted to the needs of the case.

On the other hand, there was an obvious similarity between the periodic opacity phenomenon and the periodic law of the elements. The periodic opacity phenomenon is due to a periodic variation in the properties of colloidal particles, when the number of the positive and negative ions of an electrolyte present is progressively increased. The periodic law of the elements may be defined as a periodic variation in the properties of atoms, when the number of the positive and negative elements of electricity present is progressively increased. The analogy will appear closer if, for a positive element of electricity, the term proton is used, and if, for a negative element of electricity, the term electron is used. In this way a positive ion may be defined as one which contains a relative excess of protons, and a negative ion as one which contains a relative excess of electrons. In both of the phenomena under consideration, therefore, the periodicity is associated with a progressive increase in the number of protons and electrons. Again, the periodic opacity phenomenon is more pronounced in the presence of relatively small concentrations of ions, and falls away with increasing concentrations of ions.

Similarly, atoms consisting of relatively small numbers of the elements of electricity differ from each other more than do those consisting of relatively large numbers of these elements. In both cases, therefore, the periodicity is damped in character. As the particular state of matter associated with the periodic opacity phenomenon is the colloidal one, it is evident that adsorption must play an important part, and that therefore the mechanism must depend in general upon the nature of, the arrangement of, and the spacing of, the atoms forming the surface layer of the colloid. In consonance with this the properties of atoms are, according to the Lewis-Langmuir theory (1916 (5), 1919 (6)), determined by the number and arrangement of the electrons in the surface layer, and the ease with which they are able to revert to more stable forms by giving up or taking up electrons or by sharing their outside electrons with atoms with which they combine. Again, the periodic variation in the opacity of colloids means that the total surface area of the colloid scattering light varies in a periodic manner. W. L. Bragg (1920 (7)) has shown that the diameter of atoms, and therefore the surface area also, varies in a periodic manner. It is evident, therefore, that the characteristic common to both phenomenon is that the periodicity is predominantly associated with surface area and surface constitution. In making the above comparison, mention was made of the tendency of atoms to revert to more stable forms by giving up or taking up electrons or by sharing their outside electrons with atoms with which they combine. Similarly with colloids, there will be a tendency for the particles to revert to more stable forms by giving up or taking up ions or by sharing their outside ions with particles with which they combine. In short, not only atomic matter but also aggregated matter tends to revert to more stable forms. In the electron theory the stable forms of atoms are those of the inert gases, helium, neon, argon, etc., in which the surface layer has its optimum quota of electrons. In the theory of the periodic opacity phenomenon, the stable forms of the crystallo-colloidal complex are those corresponding to the minima in the curves at which points the surface layer of the colloid has its optimum quota of ions. As an example of the application of the theory we may take the portion between the first two minima of the curve showing the effect of increasing concentration of calcium chloride on the opacity of a buffer solution of phosphates (see fig. 11). The rise in the curve represents the effect of adsorption of calcium ions with a consequent increase in the aggregation of the particles, and, therefore, in the opacity of the suspensions, and also a consequent increase in the degree of instability of the colloid. Beyond the maximum, however, any additional adsorption of calcium ions would result in the production of too great a degree of instability of the colloid. But relief from the strain produced by a highly unstable form of colloid will follow the line of least resistance, and since the ions causing the strain cannot be removed from the system, the neutralisation of that strain with the least possible expenditure of energy will be most readily brought about by the colloid-adsorbing ions of opposite charge. In the present case, therefore, chlorine ions would be adsorbed, with a consequent decrease in the aggregation of the particles, and, therefore, in the opacity of the suspension, until at the next minimum the proportions of calcium and chlorine ions would be at an optimum once more.

The research was carried out in the University of Manchester and the expenses thereof were defrayed partly by the Department of Pathology, for which I thank Prof. H. R. Dean, and partly by the Medical Research Council, to whom an acknowledgment is also due. My thanks are also due to Prof. A. V. Hill of the Physiology Department, for his crucial suggestions that experiments should be made with varying wave-lengths of incident light; to Dr. G. A. Hemsalech and Mr. H. Bell of the Physics Department, for helping to fit up part of the apparatus used for producing monochromatic light; and to my two assistants, Mr. J. W. R. Ponting and Mr. W. J. Preston, for helping to carry out the experiments.

REFERENCES.

(1) J. Holker.—"The opacity of a mixture of serum and Wassermann 'antigen' in progressively increasing concentrations of sodium chloride." 'Journ. of Path. and Bact.,' vol. 25, No. 3, 1922.

(2) J. Holker.—"The periodic opacity of Wassermann 'antigen' in progressively increasing concentrations of sodium chloride." 'Journ. Path. and Bact.,' vol. 25, No. 4, 1922.

(3) J. Holker.—"Methods of measuring the opacity of liquids." 'Biochem, Journ.,' vol. 15, No. 2. (1921).

(4) R. Bourdillon.—'Journ. Chem. Soc.,' vol. 103, p. 791 (1913).

(5) G. N. Lewis.—"The atom and the molecule." 'Journ. Amer. Chem. Soc.,' vol. 38, p. 762 (1916).

(6) I. Langmuir.—"The arrangement of electrons in atoms and molecules." 'Journ. Amer. Chem. Soc.,' vol. 41, p. 868 (1919).

(7) W. L. Bragg.—"The arrangement of atoms in crystals." 'Philosophical Magazine,' vol. 40, p. 169 (1920).

The Kinetic Energy of Electrons Emitted from a Hot Tungsten Filament.

By J. H. Jones, B.Sc., Ph.D.

(Communicated by Prof. O. W. Richardson, F.R.S. Received August 5, 1922.)

Introduction.

The first measurements of the kinetic energy of the electrons emitted from hot bodies were made by Prof. Richardson and Dr. F. C. Brown in 1907–1909.* These experiments showed that the velocity distribution among the emitted electrons was in close agreement with Maxwell's law of distribution for a gas, of molecular weight equal to that of the electrons, in thermal equilibrium at the temperature of the source. In the simple unidimensional case where the cathode and anode form parallel planes of indefinite extent, the current which flows against a retarding potential, V, depends only on the normal velocity component and with Maxwell's distribution is given by $i = i_0 e^{-2h\epsilon V}$, where $h = 1/2k\mathrm{T}$, k being Boltzmann's constant, T the absolute temperature of the source, and ϵ is the electronic charge. The mean kinetic energy of the electrons in the stream is given by the quantity $2k\mathrm{T}$. These experiments showed that with platinum, which was the only metal tried, the exponential equation was very accurately obeyed, and the average of eight determinations of k agreed with the theoretical value to within a fraction of 1 per cent. although the individual determinations differed from the average by almost ±20 per cent. These experiments were, however, subject to a number of defects, the most important of these being due to the presence of electric and magnetic fields caused by the electric currents used in heating the source.

---

* 'Phil. Mag.,' vol. 16, pp. 353, 890 (1908); vol. 18, p. 681 (1909).

Schottky* carried out some experiments in 1914, and he used the case of a filament surrounded by a concentric cylindrical anode. The effects of the magnetic and electric fields of the current used to heat the filament were avoided by an interrupted current method due to v. Baeyer. Schottky's experiments were made with carbon and tungsten, and the data were in good agreement with the requirements of Maxwell's law, except that the average energy of the emitted electrons was in every case in excess of the value calculated from the temperature of the source. This, however, was estimated from the value of the saturation current using the emission constants given by other authors. This makes his temperature determinations very uncertain, because of the known large effects on the emission of traces of certain contaminants.

In 1917, Prof. Richardson, with the assistance of Mr. Sih Ling Ting, started experiments in order to try and clear up some of these matters. In these experiments† platinum and tungsten were used as sources of emission, and in both cases abnormally high values for the average energy were obtained.

It was found, however, that the logarithmic nature of Maxwell's law was strictly obeyed. The average energy determinations differ by a factor of 2 approximately in all cases, that is, the distribution corresponded to a temperature of source of about twice as high as was actually the case.

The present work is a continuation of the above, and was carried out under the direction of Prof. Richardson, and was commenced in October, 1919. It will be shown in the sequel that the very high values of the average kinetic energy given by Ting in his paper are erroneous, and have been satisfactorily accounted for.

Theoretical Considerations.

The experiment consists in determining the current potential curves for different temperatures. From these curves it is possible to estimate the average energy of the stream of electrons. We require to find a relation between the negative current passing from a filament to a concentric cylindrical anode, and the potential difference applied across this gap. The number of electrons which have a radial velocity between z and $z+dz$ is given by the expression*

---

* 'Ann. der Physik,' vol. 44, p. 1011 (1914).

† 'Roy. Soc. Proc.,' A, vol. 98, p. 374.

$$N_z\,dz = 4N\,(h^3m^3/\pi)^{\frac{1}{2}} z^2 e^{-hmz^2}\,dz. \tag{1}$$

The number of electrons which have a velocity between w and $w+dw$ in a direction parallel to that of the filament is given by*

$$N_w\,dw = N\,(hm/\pi)^{\frac{1}{2}}\,e^{-hmw^2}\,dw. \tag{2}$$

Now suppose V is the applied potential difference between the filament and the cylindrical anode, then all electrons which have a velocity z' greater than $(2\epsilon V/m)^{\frac{1}{2}}$ will reach the cylinder, whilst all electrons with a smaller velocity than this will be turned back to the filament and will not contribute to the electron current. The current at this potential difference will be given by

$$i = \epsilon \int_{z}^{+\infty} N_z\,dz \int_{-\infty}^{+\infty} N_w\,dw. \tag{3}$$

Substituting the values of $N_z dz$ and $N_w dw$ from (1) and (2) in (3), and putting $N\epsilon = i_0$ (the current at zero potential difference), we get

$$i = 2\pi^{-\frac{1}{2}} i_0 \left\{ e^{-n}\sqrt{n} + \int_{\sqrt{n}}^{+\infty} e^{-x^2}\,dx \right\}, \tag{4}$$

where $n = 2h\epsilon V$.

It is assumed in the calculation that the radius of the filament is negligible compared to that of the cylinder.† We have also assumed that the cylindrical electrode is infinitely long compared to the length of the filament. For this reason the limits of w have been taken to be $\pm\infty$. If now we plot n against $\log i/i_0$ we shall get a curve which is a straight line for all values of $n > 3$, and the slope of this straight portion is $(\log i/i_0)/n = 0{\cdot}405$. This curve is graphed in fig. 9.

The average kinetic energy of the stream of electrons‡ is

$$1/h = 2kT, = 0{\cdot}405\left(\frac{2\epsilon V}{\log i/i_0}\right).$$

These $(\log i/i_0, V)$ curves are obtained at different temperatures of the filament and their slope gives us a measure of the mean kinetic energy.

Effect of Space Charge.

At high temperatures, when the current density become large, the effects due to the mutual repulsion of the electrons cannot be neglected, and the theory given only applies to those cases where the concentration of the electrons in the space is very small.

---

* O. W. Richardson, 'Emission of Electr. from Hot Bodies,' 2nd ed., p. 179.

† *Cf.*, however, Schottky, 'Ann. der Physik,' vol. 44 (1914).

‡ O. W. Richardson, 'Phil. Mag.,' vol. 16, p. 353 (1908).

We can see, in a general way, the effect of this space density of the electrons on the current potential curves. When the concentration is small (such as with currents of the order of 10^{-9} ampères) the electric force applied between the electrodes will practically be the only force acting on the electrons, but for larger concentrations there will be an additional force due to the mutual repulsion of the electrons. Further, the amount of this force will vary with the concentration, and consequently with the applied potential difference. The effect of this will be not only to shift the curves bodily along the potential axis, but also it will tend to change the slope. Suppose A, B, fig. 1, represents the hot electrode, and C, D, the cold, and let the applied potential difference be V_0 represented by E, F. When the concentration in the space between A, B, and C, D is very small, the potential distribution will be represented by the straight line O, F. As the currents (and concentration) increase the linear distribution will no longer hold, and we get a distribution something like O, G, F, in which there is a point of minimum potential at G. It is this minimum potential which will control the current to the electrode C, D, and this potential difference V′ must replace V in the equation

$$i = i_0\, 2/\sqrt{\pi}\left\{e^{-u}\sqrt{u}+\int_{\sqrt{u}}^{+\infty} e^{-x^2}\,dx\right\},$$

so that we must know the value of V′ when we are working with high density electron currents. This minimum potential set up in the space will be more important at the lower potential differences when the electron currents are large. As the current to the anode is decreased by

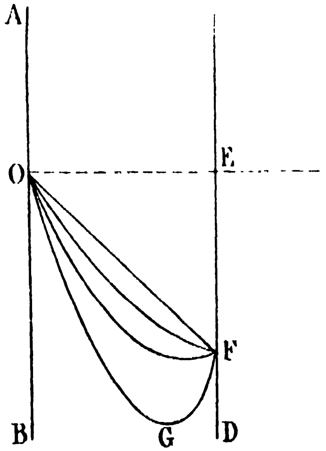


Fig. 1.

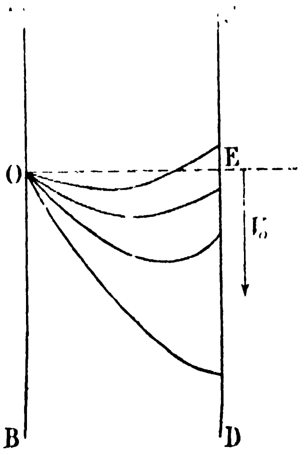


Fig. 2.

increasing the potential difference between the electrodes, the value of $V' - V_0$ will continually decrease. The potential distribution curves will be something similar to those shown in fig. 2. As the potential difference V_0 is increased the point of minimum potential will shift towards C, D, and finally coincides with C, D, when V_0 is sufficiently large. Let us now examine the effect of this on the current-potential curves. Suppose A, B, C, fig. 3, represents the curve in which the current to C, D is plotted against V_0 the potential difference registered on the voltmeter. The curve in which the current is plotted against V', which is what we require, will be something similar to A′, B′, C′, fig. 3, where the shift B, B′ will be equal to $V' - V_0$. It is clear that this curve must be steeper than A, B, C′, so that the values of

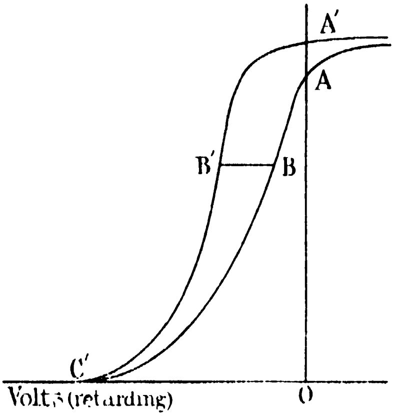


Fig. 3.

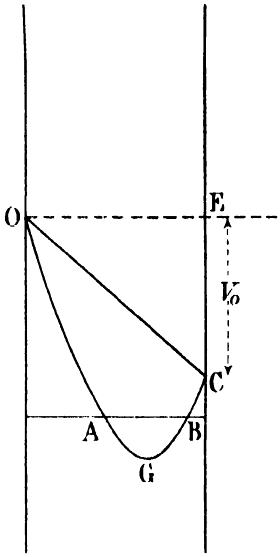


Fig. 4.

$(\log i/i_0)/V$ will be lower than the true values. It is, however, very difficult to make the necessary correction in the case of a filament and cylindrical anode.

The following calculations show the difficulty in evaluating this correction. We shall, first of all, deal with the case of two parallel plane electrodes:—

Let us calculate the density of the electrons at a point where the potential is V (negative) with respect to the hot cathode. We assume also that the potential distribution takes the form shown in fig. 4. V_0 is the applied potential difference, and we suppose that a minimum potential V' is set up in the space between the electrodes. The current to the anode is determined by the value of this minimum potential V'. G is the point of this minimum potential. Consider first the density at a point A (potential difference = V) in the portion O, A, G of the curve. The density at this point is :—

$$\rho_A = \epsilon \int_a^{+\infty} N_u \, du / \sqrt{(u^2 - a^2)} \int_{-\infty}^{+\infty} N_v \, dv \int_{-\infty}^{+\infty} N_w \, dw +$$

$$\epsilon \int_a^{a'} N_u \, du / \sqrt{(u^2 - a^2)} \int_{-\infty}^{+\infty} N_v \, dv \int_{-\infty}^{+\infty} N_w \, dw,$$

where the limits a and a' are given by

$$a = (2\epsilon V/m)^{\frac{1}{2}}, \quad a' = (2\epsilon V'/m)^{\frac{1}{2}}.$$

Putting in the values of $N_u \, du$, $N_v \, dv$, and $N_w \, dw$, given by Maxwell's distribution law, we get

$$\rho_A = i_0 \sqrt{(\pi h m)} \, e^{-aV} \left\{ 1 + 2/\sqrt{\pi} \int_0^{a(V'-V)} e^{-x^2} \, dx \right\}.$$

Similarly, for a point B in the portion GBC, we get

$$\rho_B = i_0 \sqrt{(\pi h m)} \, e^{-aV} \left\{ 1 - 2/\sqrt{\pi} \int_0^{a(V'-V)} e^{-x^2} \, dx \right\}.$$

Applying Poisson's equation, and expressing the error function as a series, the first integration can be performed.

The second integration cannot, however, be elevated in a simple way. A graphical solution has been given by Fry.* In the case of a filament and concentric cylindrical anode the differential equations are even more difficult to solve.

Suppose the radius of the filament is r_0, and that of the cylinder is r_1, then the density of the charge at any radius r can be written down as

$$\rho = \epsilon r_0 / r \left\{ \int_a^{+\infty} N_z \, dz / \sqrt{(z^2 - a^2)} \int_{-\infty}^{+\infty} N_w \, dw + \int_a^{a'} N_z \, dz / \sqrt{(z^2 - a^2)} \int_{-\infty}^{+\infty} N_w \, dw \right\}.$$

Here $N_z dz$ is the number of electrons which have a radial velocity lying between z and $z + dz$, and a and a' have the same meaning as before.

Poisson's equation for this case is :—

$$d/dr \, (r \, dV/dv) = -4\pi\rho r.$$

The right-hand side of this equation is a complicated function of V, the potential, and involves the error function. It has not been found possible to obtain a solution.

Two special cases of this problem have been solved: (1) by Langmuir,† where the initial velocities of the electrons were neglected; and (2) by

---

* Fry, 'Phys. Rev.,' p. 441 (April, 1921).

† 'Phys. Rev.,' vol. 2, p. 453 (1913).

Laue,[*] who dealt with the case of an electron atmosphere in equilibrium with the hot body.

Method of Experiment.

The tube consisted of a tungsten filament, generally about 6 cm. long, supported by two thick copper leads and surrounded by a co-axial copper cylinder enclosed in a glass tube. During the earlier stages of the work, the method of procedure was to exhaust the tube in a vacuum furnace, by a Gaede mercury pump. The furnace was kept at about 750° K., and the pumping carried on for about 30 hours, until the occluded gases in the metallic parts and in the walls of the tube were driven out. Liquid air was then applied to a charcoal condenser placed in series with the tube and Gaede pump. The liquid air is kept round the charcoal for about 10 hours, and during this time the furnace was allowed to cool, and opened so that the filament could be glowed out by a strong current for a quarter of an hour or so.

Later tubes were kept permanently on the pumps, and a mercury vapour diffusion pump was used in conjunction with the Gaede. In this way very high vacua were obtained, and in most cases the pressure was measured by means of a Knudsen gauge specially constructed for this purpose. The gauge used was similar to that described by Dushman and Found.[†] The particular gauge used was so designed that pressures of the order of 10^{-8} mm. of Hg could be measured.

The diffusion pump was first of all tried with one liquid air trap on the high vacuum side, but only vacua of the order of 10^{-6} mm. of Hg could be obtained. A second trap was then introduced, as it was thought that some mercury vapour passed through the first trap and so kept the vacuum down. This was found to be the case, and the vacuum was very much improved with the second trap. In fact, the pressure was then too low to be measured by the Knudsen gauge, and the vane only showed an appreciable deflection when the strip was made red hot. This meant that the pressure was less than 10^{-8} mm. of Hg. The chief source of trouble in this experiment is due to the fact that tungsten is very active, and readily attacks nearly all gases when it is heated, and is extremely sensitive to small traces of water vapour in the tube. It is known that glass gives out an unlimited supply of water vapour, and this attacks the hot tungsten, forming probably an oxide layer on the surface, which cuts down the emission enormously. The action of some condensible vapour on tungsten is shown very clearly by the following test which was made. An additional liquid air trap in the form of a U-tube was introduced as close as possible to the tube. When liquid air was applied to this, the thermionic emission from tungsten was considerably increased. The liquid air was then removed and the emission immediately decreased, and in some cases the filament ceased to emit altogether. The emission could be started again by heating the tube in the vacuum furnace as before. That there was a change in the condition of the filament was clearly shown by the difference in the heating current required to bring the filament to a given resistance. The larger the thermionic emission the larger was the current necessary to keep the filament at a given resistance. These changes in the heating current are far too large to be accounted for by the extra cooling due to the emission. They are probably caused by changes in the condition of the surface of the filament, either causing changes in the radiating power or else changing the specific resistance. When the thermionic emission is greatest, we would expect the surface of the filament to be cleanest. If the variations in the filament current depend on the change in the radiating power of the surface, then the pure metal surface should radiate better than the contaminated surface. The reverse would be expected to be the case. On the other hand, it is unlikely that the resistance would be altered very much by these surface films. This point is now being investigated with the aid of an optical pyrometer.

---

[*] 'Jahr. der Radio. und Elektr.' (Nov., 1918).

[†] 'Phys. Rev.,' p. 7 (Jan., 1921).

The Method of Measuring the Kinetic Energy of the Electrons.

The filament generally forms one arm of a Wheatstone bridge with a battery for supplying the heating current. Two other arms are taken from a Post Office box, and the remaining arm is a manganin resistance comparable to the resistance of the filament.

Two make-and-break switches are inserted, one in the heating circuit and the other in the thermionic circuit, with their phases so arranged that the thermionic current is measured only when the heating current is off. The method used by Ting in his experiments is shown in fig. 5. K represents the commutator with three terminals which serves the purpose of two switches mentioned above. When the commutator is rotating, the middle terminal makes contact alternately with the two side ones. P, Q, R, S represents the Wheatstone bridge. The two resistances, P and Q, are taken from a Post Office box and are of the order of 2000 ohms. S is a resistance of manganin wire of the same order as the filament R.

The potential difference between the cylinder and filament is supplied by the battery, B′, and rheostat, P, and is measured by the voltmeter, V. The thermionic current is measured by the galvanometer, G, connected with a shunt, L. The commutator, K, is set rotating at such a speed that the temperature does not drop appreciably during the interval the heating current is cut off. It is essential to have good contact between the brushes and the commutator, otherwise the temperature of the filament will be fluctuating. With a little practice, however, this can be easily done.

This method, as it stands, is not correct and gives wrong values for the mean energy of the electrons. If the heating current be reversed through the filament, then not only is the magnitude of the electron current different, but entirely different current-potential curves can be obtained. Passing the current in one direction, we get curves which are logarithmic, but the apparent mean energy of the electrons will be very high, whilst if the current is passed in the other direction, the curves are not even logarithmic. The logarithmic curves are obtained when the direction of the heating current is such as to produce a positive drop of potential from the filament to the commutator, *i.e.*, the filament will be positive with respect to the commutator (and cylinder). These are the curves which were probably obtained by Ting. They do not, however, represent the true electron emission from the filament, as will be shown below.

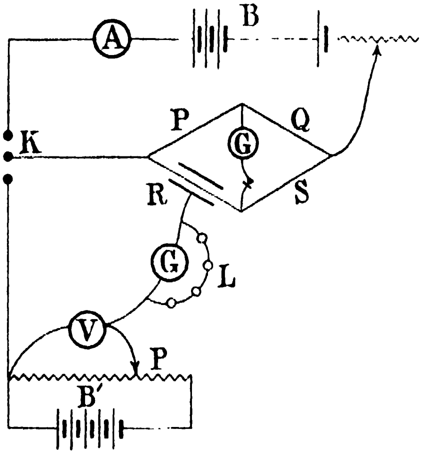


FIG. 5.

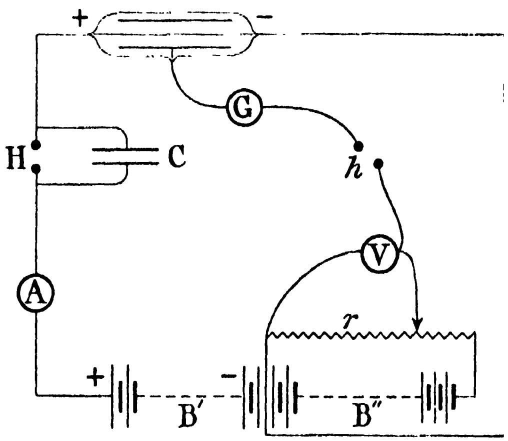


FIG. 6.

With the potential drop in the other direction, *i.e.*, filament negative with respect to commutator (and cylinder), the current-potential curves are not logarithmic, and in most cases a potential of about 15 volts is required to stop the current to the cylinder, whereas, in the other case, the current could be stopped by a potential of 3 volts or so. In fig. 10 (*a*) and 10 (*b*) are shown curves obtained in this way. The curves in fig. 10 (*a*) are similar to those obtained by Ting. It is seen that these curves are logarithmic, but that the slope of the log-curve is about half the slope of the theoretical curve, assuming Maxwell's distribution law to hold. This was the conclusion arrived at by Ting from his experiments.

Fig. 10 (*b*) shows the type of curves obtained when the heating current flows in the reverse direction. It is observed that these curves are not logarithmic.

Schottky's Method.

The method employed by Schottky in his experiments in 1914 is represented in fig. 6. In this diagram H and *h* are the two switches. A condenser, C, was put across the switch, H, in the heating circuit, though in his paper Schottky gives no reasons for introducing this condenser. However, he seems to have avoided the more serious instrumental errors, as his results show much closer agreement with theory than those of Ting.

Let us now consider the complications which were likely to have occurred in Ting's experiments. In the first place, complications will arise due to the charging up of the cylinder whilst the heating current passes through the filament and the thermionic circuit is kept open. When this circuit is closed there will be a discharge through the circuit, and as the commutator rotates at a high speed, this is likely to have a considerable effect. Further, there is a considerable potential drop along the filament when it is being heated by the current, and as this potential drop is varied rapidly by the commutator, it will cause capacity currents through the thermionic circuit. If there are n breaks per second, and if the change in potential of the filament is V due to each break, and C is the capacity between the filament and the cylinder, etc., the current set up in this way will be $i = nCV$. The capacity, C, will generally be very small, but the quantities n and V are considerable, so that these capacity currents will not always be negligible, especially when the electron currents are themselves small, either when the temperatures are low, or when there are large applied potential differences. For example, consider a filament about 10 cm. long surrounded by a co-axial cylinder of 1 cm. radius. Let the radius of the filament be 0·005 cm., as was the case in some of Ting's experiments. The number of breaks (n) is 200/sec., and V will be about 10 volts.

The capacity (of cylinder and filament) $= e/(2 \log_e b/a)$ where $e = 10$ cm., $b = 1$ cm., $a = 0{\cdot}005$ cm.

This works out to be a capacity of 1 cm. (roughly). The current will be $i = 10^{-9}$ ampères.

We have only taken into account the capacity of the filament and cylinder, and it is quite possible that the current will be bigger than this value.

Two different methods have been devised which appear to overcome all these instrumental difficulties. The problem was to eliminate the charging up of the cylinder due to the flow of the electron current whilst the heating current was flowing, and also due to the capacity effects. Figs. 7 and 8 show these two methods. The first method tried is shown in fig. 7. This differs from the methods already discussed in the manner in which the commutator is used. Here only two brushes are used and the thermionic circuit is permanently closed. A resistance, M, is introduced between the commutator, K, and the bridge, and the heating current is passed in the direction which makes a positive potential drop between the filament and the commutator, K. If there is no applied potential difference, that is, the voltmeter reads zero, then when the current is passing through the filament, the cylinder, which will be at the same potential as the commutator, will be at a high negative potential with respect to the filament, and this potential difference is arranged to be, by means of M, enough to stop the electron flow altogether. As soon as the heating current is off, the filament and cylinder are at the same potential (except when there is an applied potential difference as registered by the voltmeter). It is very important to make M a non-inductive resistance because of the rapid make and break of the circuit. It is clear that this method gets rid of the complications discussed above. The charging up of the cylinder whilst the heating current is flowing will be stopped, and also the capacity effects will be nullified. This is the case because the capacity current at " make " will be balanced by the reverse current at " break." Some of the results given below have been obtained in this way.

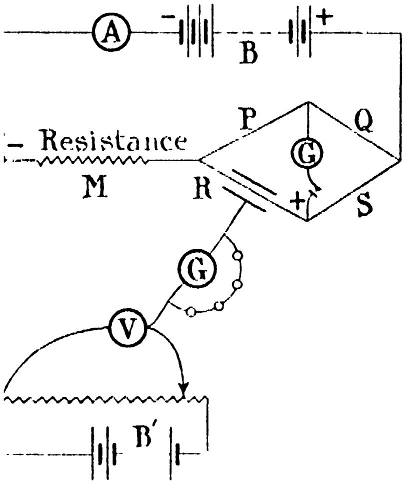


Fig. 7.

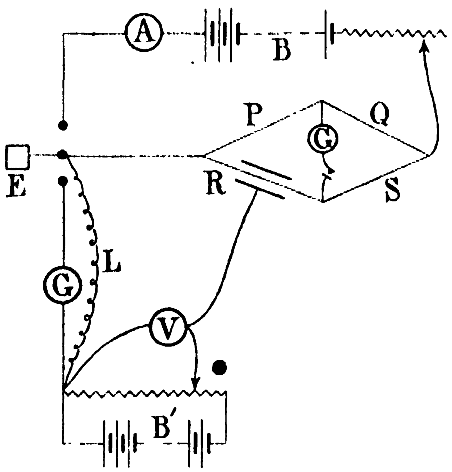


Fig. 8.

Another method, and this is the best, was used during the latter part of the work. This is shown in fig. 8. The disadvantage of the above method is that the heating current can only be passed in one direction through the filament. This, however, does not apply to the method shown in fig. 8.

Here three brushes are again used, but the electron circuit is kept closed by means of the high resistance, M, which serves for another purpose as well. It will be observed that the galvanometer for measuring the electron current has been shifted, and is now so arranged that it will only measure the electron current whilst the filament heating current is cut off. The electron current, which passes to the cylinder during the heating of the filament, has now a circuit through the high resistance, and will not charge up the cylinder. This resistance can be made to serve as a shunt for the galvanometer, G_1. It has been found practicable to earth the middle brush of the commutator. The current potential curves obtained in this way are independent of the direction of the heating current; further, the thermionic current has the same magnitude in both directions, which could not be obtained by the other methods. Identical results have been obtained with these two methods, *i.e.*, methods of figs. 7 and 8.

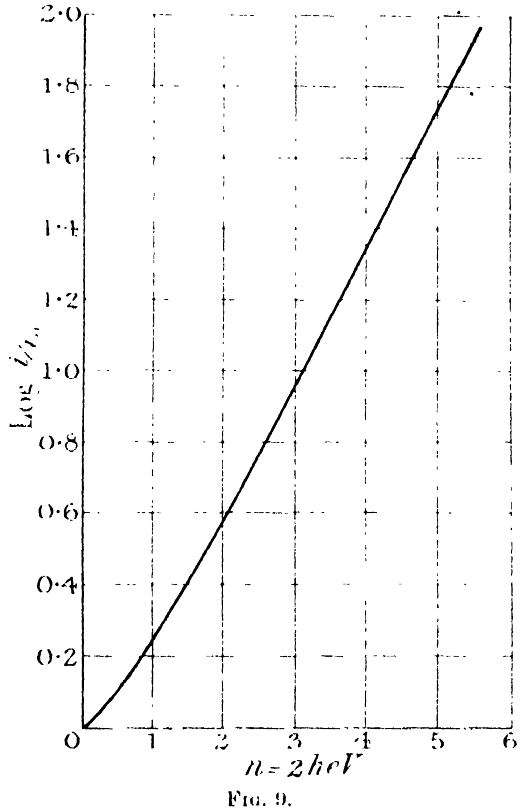


Fig. 9.

Estimation of the Temperature of the Filament.

The temperature of the filament was deduced from its change of resistance, making use of data given by Langmuir on the change of resistance of tungsten with temperature.*

Results of the Experiments.

First Tube.—Tungsten filament and copper cylinder. Length of filament, 6·5 cm. Length of cylinder, 10·5 cm. Diameter of filament, 0·005 cm. Diameter of cylinder, 2 cm.

The Tables below give the values of the slope of the log curves and the temperature of the filament. The theoretical values for the slope, assuming Maxwell's distribution law to hold, are also given for comparison. A current potential and the corresponding log curve are shown in fig. 11.

Table I.

Filament temperature.	$\frac{\text{Log } i/i_0}{\text{V}}$.		Error.
	Experimental.	Theoretical.	
° K.			per cent.
1640	3·10	2·87	+8
	2·75	2·87	−4
	2·68	2·87	−7
1770	2·65	2·66	0
	2·50	2·66	−6
	2·60	2·66	−2
	2·60	2·66	−2
1800	2·63	2·61	+1
	2·55	2·61	−2
1920	2·30	2·45	−6
	2·00	2·45	−18
	2·15	2·45	−12
1980	2·20	2·38	−8

---

* 'Phil. Mag.,' vol. 7, No. 3 (March, 1916).

Table II.—Second Tube.

(This tube was exactly similar to the first.)

Filament temperature	$\frac{\text{Log } i/i_0}{\text{V}}$		Error
	Experimental.	Theoretical.	
K.			per cent.
1650	2·43	2·85	−15
	2·40	2·85	−16
1850	2·75	2·55	+8
	2·25	2·55	−12
	2·35	2·55	−8

Table III.—Sealed-off Tubes.

(Dimensions as before.)

Filament temperature.	$\frac{\text{Log } i/i_0}{\text{V}}$.		Error.
	Experimental.	Theoretical.	
° K.			per cent.
1500	3·23	3·14	+3
	3·00	3·14	−4
	3·00	3·14	−4
1600	2·95	2·94	0
	3·25	2·94	+10
	3·00	2·94	+2
1730	2·65	2·72	−2
	2·55	2·72	−6

Discussion of Results.

Tables I, II, and III above give the values of the slopes of the $(\log i/i_0)/\text{V}$ curves at different temperatures.

Those in I and II were obtained with tubes whilst on the pumps, and the pressure in almost every case was of the order of 10^{-8} mm. of Hg. The method used was that shown in fig. 7. Those given in Table III were obtained with tubes which had been sealed off from the pumps after having been exhausted as low as possible. The method shown in fig. 8 was used.

More attention has been paid to the lower temperatures because of the space charge effects at the higher temperatures.

Referring back to fig. 9, where the theoretical curve is shown, it is seen that it only approximates to a straight line when $n > 3$, *i.e.*, $2h\epsilon V > 3$. This

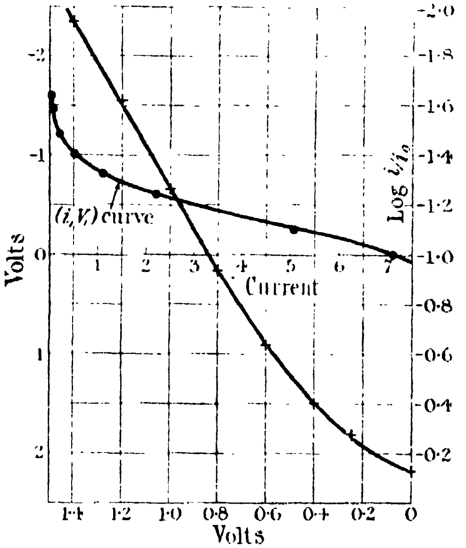


FIG. 10*a*.

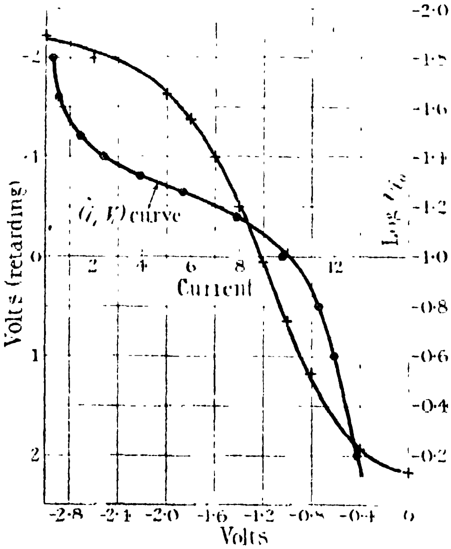


FIG. 10*b*.

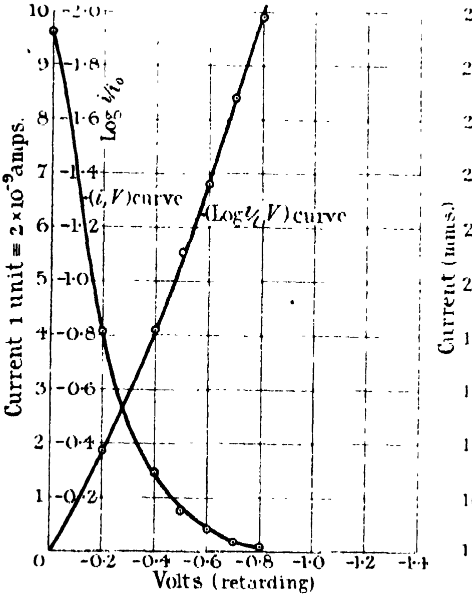


FIG. 11.

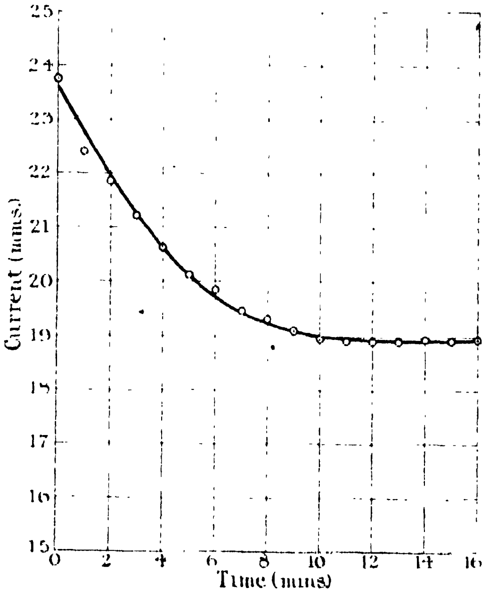


FIG. 12.

means that for a temperature of 1600° K. or so V must exceed 0·4 volt before we get the linear relation which is the important part of the curve.

As the temperature increases this critical value for V also increases. At these retarding voltages the electron current is already cut down to a small fraction of its saturation value, and this means that the current in the important part of the curve will be very small, and consequently makes the observations more difficult to take and liable to bigger percentage errors. This is one disadvantage of the cylindrical case as compared to the unidimensional method, where the curves are logarithmic throughout their range. There are, on the other hand, many disadvantages to the parallel plate method, the most serious of these being the difficulty of getting (1) a truly plane surface, and (2) a uniform temperature.

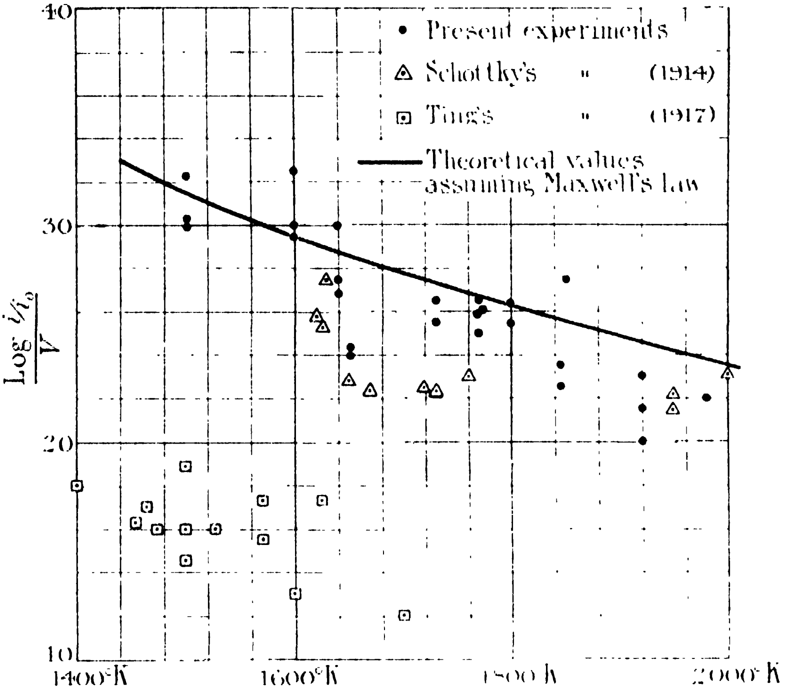


FIG. 13.

It will be observed from Tables I, II, and III (or fig. 13) that the tendency is for the values of $(\log i/i_0)/V$ to be smaller than the theoretical values assuming Maxwell's law to hold. This is also the case in Schottky's results, and, in fact, all of his values are below the theoretical curve. The effects mentioned above, due to the mutual repulsions of the electrons, would account for these deviations at higher temperatures when the currents are large, but at the lower temperatures these variations must be due to another cause. There are strong reasons for believing that these disagreements result from the formation of an oxide film on the surface of the filament, due to minute traces of water vapour which appears to be given off indefinitely from the glass walls of the apparatus. The presence of these films is shown by the changes in the thermionic emission accompanied by changes in the heating current.

The kind of variation that occurred is shown in the Table below:—

Heating current.	Thermionic current.
1·54 amps.	$21·6 \times 10^{-9}$ amps.
1·50 "	12·9 "
1·40 "	8·6 "
1·30 "	2·5 "

The increments in the heating current are too large to be accounted for by the small extra cooling effect due to increase in the thermionic emission, and have not yet been satisfactorily explained.

When the vacuum is very good (about 10^{-8} mm. of Hg) these changes will occur suddenly in either direction, and it makes the thermionic current very unstable. If, however, the filament is kept hot for a few hours in a vacuum of this order, the emission will again become stable, and will be considerably increased. Similar effects have been observed by K. K. Smith.*

It is very probable that the discrepancies in the values of the mean kinetic energy are in some way connected with the degree of purity of the surface. The figures given were nearly all obtained after the filament had been previously heated to a high temperature for an hour or so and the emission had increased to a maximum, as it was found that the best results were always obtained in this way. If the filament was not previously heated to a high temperature, the emission would generally decay with time. One of these decay curves is shown in fig. 12. The emission at first dies down rapidly with the time, but the rate of decay gradually decreases, and eventually the current will assume a constant value. If the filament is allowed to cool down and then heated up again, the same process will be repeated. This can be explained by the formation of a compound on the surface which is unstable at low temperatures. Similar time variations have been observed with the emission of positive ions from hot bodies. The results given in Table III have been obtained with sealed-off tubes, and are, on the whole, more satisfactory. If the tubes had been properly exhausted, and special precautions taken to keep away traces of water vapour before sealing off, then the emission would be very steady and the current potential curves show good agreement with theory. These conditions could only be maintained for a few days after sealing off, and the emission would gradually become smaller, and begin to show the time variations described above. The current potential curves would now be less steep, and may be as much as 20 per cent. below the initial value. It may be mentioned here that if the pressure of gas in the tube is of the order of 10^{-3} or 10^{-4} mm. of Hg, then these low-slope curves will almost invariably be obtained.

---

* 'Phil. Mag.,' vol. 29, p. 821 (1915).

The values of $(\log i/i_0)/V$ given in Tables I, II, and III are plotted against the temperature of the filament in fig. 13. The theoretical curve is given on the same diagram. The agreement is, on the whole, fairly good when we take into consideration the various sources of error in the experiment. Taking the mean of all the figures in Tables I, II, and III we get an agreement to within 4 per cent. of the Maxwell value, whereas if we take the mean of the figures in Table III alone, we get an agreement to $\frac{1}{3}$th per cent.

Conclusions.

1. When allowance is made for experimental and secondary effects the distribution of energy is found to agree with that given by Maxwell's law.
2. Of the experimental errors the most serious are probably due to the difficulty of measuring accurately the small currents involved and the temperatures. These lead to uncertainties, which in individual experiments may amount to as much as 10 per cent.
3. The secondary effects probably arise from contamination of the heated surfaces. This tends to increase the apparent energy of the electrons emitted and this increase may amount to as much as 20 per cent.
4. The abnormal electron energies found by Ting, which were as much as 100 per cent. in excess of the Maxwell distribution value, do not appear under satisfactory experimental conditions.

I wish to express my best thanks to Prof. O. W. Richardson for his valuable advice during the course of the work.

The Production of Artificial Vowel Sounds.

By Sir R. A. S. Paget, Bart.

(Communicated by Sir W. Bragg, F.R.S. Received October 18, 1922.)

Observations by ear made by the writer indicated that each of the vowel sounds in his own voice consisted, when breathed (*i.e.* without phonation), of two component notes due to resonances in the oral cavity.*

The resonances observed—which are given to the nearest semitone—are set out in the accompanying Table or Chart, and a description of the observations which it summarises forms the subject of a separate communication to the 'Journal of the International Phonetics Association.'†

Experiment was made with a view to the synthetic production of vowel sounds by a resonator so shaped as to give similar resonance to those observed in the voice. It being understood that the experiments of previous investigators had not proved conclusive, either as to the nature or the reproduction of vowel sounds by resonance, the present experiments were carried out *de novo* without reference to the results of prior research.

A resonator (No. 1) was made in plasticine in rough imitation of the form and dimensions of the human mouth-cavity (see fig. 1, *a* and *b*), but in which a closed bulbous cavity took the place of the pharynx. This was found, on tapping the model or blowing across its mouth, to give two resonances, 1824∼ and 406∼. The model was progressively altered by raising the tongue, reducing the projection of the lips, and enlarging the mouth, and the alteration of resonances were noted at each modification.

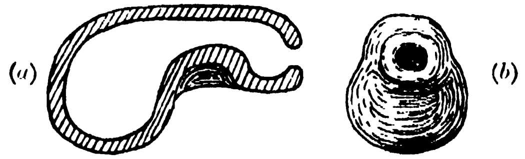


Fig. 1 (*a*) and (*b*).—Original ɐ (earth) model, No. 1 (without larynx): (*a*) Section, (*b*) Front elevation.

The model having been tuned to 2434/362 gave a clear breathed i (as in eat) by blowing across its mouth.

---

* Letter of March 3, 1922. 'Nature,' No. 2733, Vol. 109, p. 341.

† 'Vowel Resonances,' by Sir Richard Paget, Bart., International Phonetic Association, Daniel Jones, University College, London, W.C. 1. 1922. Price 2*s*.

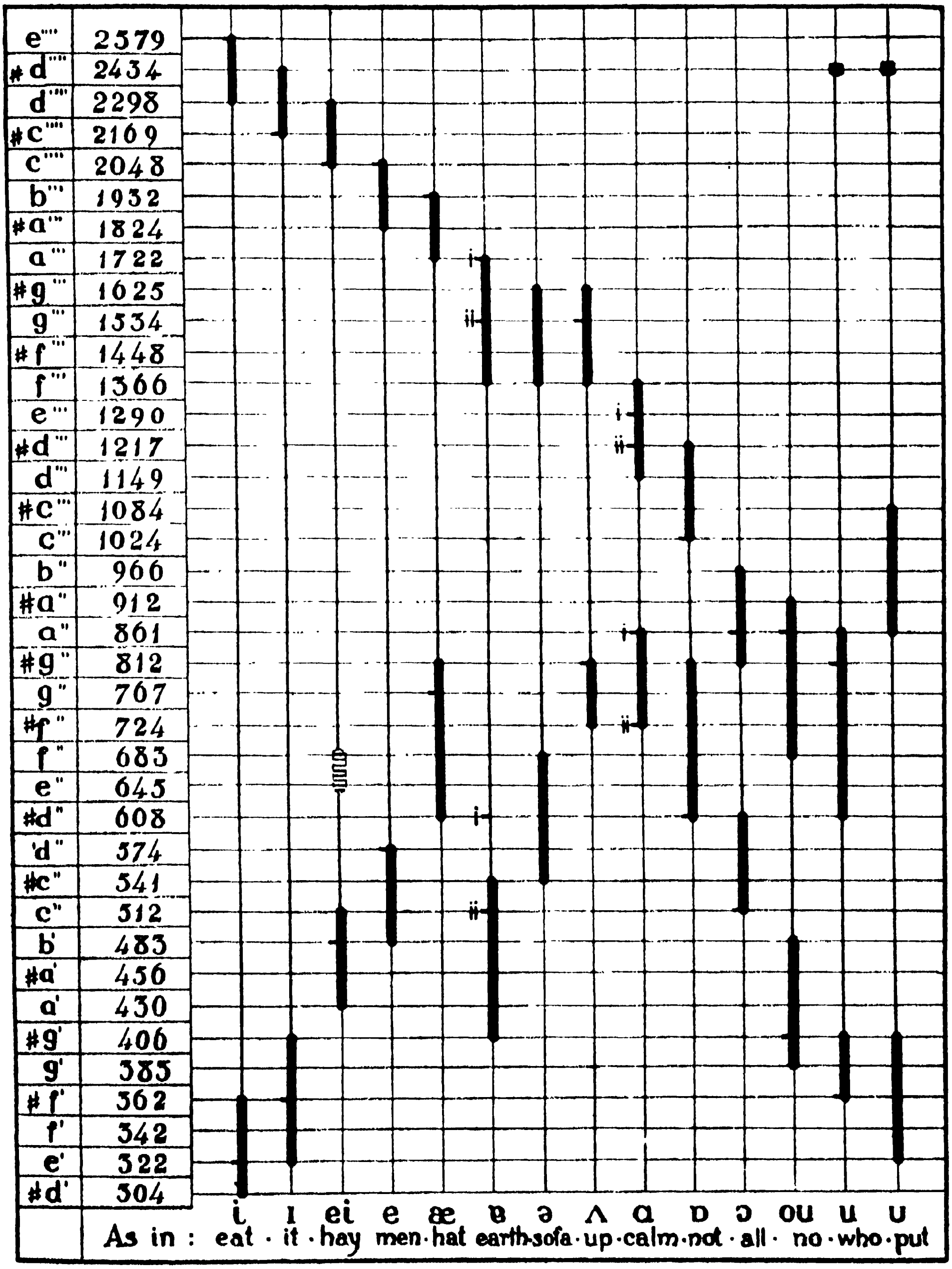


Table or Chart of vowel resonances.

The cavity having been thus tuned to give a breathed vowel, a small rear orifice was made and connected to an artificial larynx formed of a rubber strip fixed edge-wise across a flattened air passage formed in paraffin wax (see fig. 2) and encased in plasticine (see fig. 3). The larynx when blown by mouth gave a note of about g (192~) which could be raised progressively about an octave by increasing the air pressure. When this larynx was added (which altered the resonance to 2434/456) the voiced sound given by the model was not i but a muffled ɐ (as in earth).

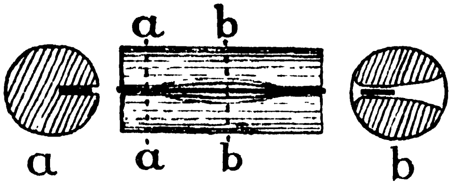


Fig. 2.

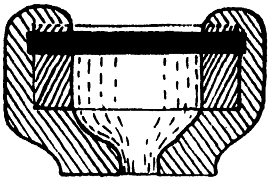


Fig. 3.

Fig. 2.—Larynx in plan and transverse sections (without plasticine casing).
Fig. 3.—Larynx in longitudinal section.

A similar model, No. 2, was made with a larger mouth and a lower position of the tongue. With this model the upper resonance could not be brought below 1722, but at 1722/724 it gave a breathed æ (as in hat) which was improved by tuning to 1824/608. A larynx was added and the model then gave a voiced æ heard as uæ uæ when the mouth was covered and uncovered with the fingers.

The model was again progressively altered (by trial and error), and at 1722/608,—*i.e.* lower resonance as before, upper resonance one semitone lower—it gave a recognisable ɐ (earth), though the upper resonance was one semitone and the lower two semitones higher than those heard in the voice. The position of the larynx was altered from the underside to the back of the model without alteration in the vowel character (see fig. 4).

A model No. 3 (see fig. 5) was made similar to the above, but with flat sides and a mouth 35 × 14 mm., giving with larynx 1024/683, and a sound between

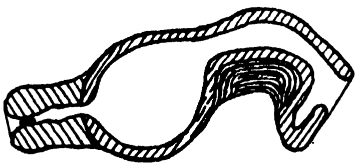


Fig. 4.

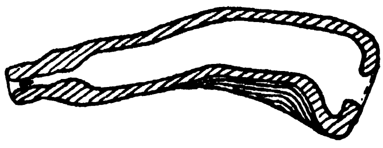


Fig. 5.

Fig. 4.—Original æ (hat) model altered to ɐ (earth) and with larynx added.
Fig. 5.—Model No. 3—ɑ (calm) (final shape).

æ (hat) and ə (as in so*fa*). This model was progressively modified and at 1217/812 gave an appreciable "ɑ" sound (as in calm).*

The model was further modified to a more convenient shape by bringing the back resonator into approximate alignment with the front; it then gave 1290/812. The upper note was heard only on tapping the *back* portion of the model. It is probable (though not recognised at the time) that the *upper* resonance in this case was produced at the back of the model (see Models Nos. 9, 13 and 16, figs. 12, 16, and 19).

Experiments on Tuning.—Over 50 experiments were made with plasticine models to test the resonance effects of modifications of the form of the cavity, the size of opening and projection of "lips," the two resonances being noted at each modification.

It was found that enlarging the mouth raised the lower resonance; it sometimes raised and sometimes lowered the upper resonance. Increasing the projection of the lips lowered both resonances. Raising the front of the tongue upwards and forwards raised the upper resonance but lowered the lower resonances. Pressing the back of the tongue back, or reducing the capacity of the back cavity, counteracted the effects of a forward and upward movement of the front of the tongue.

Thus it is possible to produce the same pair of resonances with the tongue pressed upwards and forwards and the back cavity reduced or with the tongue low and the back cavity enlarged. In the former case, however, the resonating cavities are smaller and the vocal resonance diminished, though the vowel character is unchanged. This effect has an obvious bearing on voice production.

In view of the failure of No. 1 to give a voiced i, a further model was made (No. 4) from a cardboard tube closed with plasticine (see fig. 6), and tuned to 2434 and 322, with the object of increasing the whistle action of the front cavity, which appeared to be deficient. This model gave a recognisable breathed i (as in eat), and a sibilant voiced i when a larynx was added.

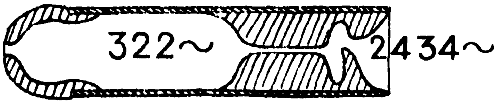


FIG. 6.—Model No. 4—Cardboard tube i (eat) model.

u (as in who).—In view of the whistle-like character of the upper resonance in a breathed u (as in who), a model (No. 5) was made with a whistling front cavity tuned to 683, a back cavity giving 215, and with a larynx attached to the back cavity (see fig. 7).

---

* The upper resonance 1024, etc., were originally recorded as an octave higher.

This model, when blown without a larynx, gave a recognisable breathed u, and a fair though sibilant u when softly blown with larynx, but when strongly blown the u character was wholly lost.

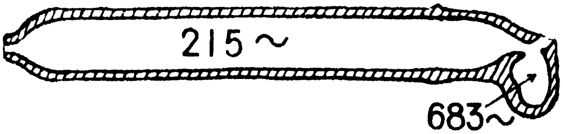


Fig. 7.—Model No. 5—u (who) front whistle model.

To test the possibility of inducing resonance without true whistle action a resonator, No. 6, was made about 70 × 40 mm. in cross-section with an orifice 50 × 30 mm., as in fig. 8, and when blown without larynx was found to give a well defined "whistle note"—the disturbance due to the entry of the jet of air being sufficient to cause resonance in the cavity.

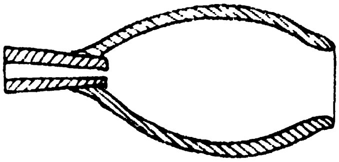


Fig. 8.—Model No. 6—Single resonator with jet inlet.

Experiment was made with this model by modifying the capacity and mouth orifice to obtain u by a single resonator, but without success.

Added Resonance.

It was found that if the hands be clasped, so as to form a cavity (see fig. 9) (open fore and aft) with a resonance of 1024, and held against the lips while sounding the various vowels,

a (calm) and ɒ (not) became like ɔ (all),
ou (no) and ɐ (earth) became like u (who).

A plasticine resonator (No. 7) of the same resonance (1024) and with front orifice 10 × 12 mm. was made (see fig. 10) and tried with similar effect; the additional resonator was then attached to the front orifice of the ɒ model. This gave a fair voiced u 812/342. When attached to the ɑ model a fairly good o 1084/383 was produced.

It was found that a good u could also be produced by blowing with the mouth into an artificial larynx attached to a single resonator (No. 8) (see fig. 11), 50 mm. diam. × 100 mm., with front orifice 1 cm., but that the u character was lost *if the shape of the operator's mouth was altered during blowing.* The model was blown with bellows and gave little or no u.

An additional plasticine cavity was added *behind* the larynx and blown with bellows, and the u character was restored. A fairly good u was also obtained by adding an additional resonator (912 with both orifices open, front orifice 16×18 mm.) in front of the single resonator (No. 8). The vowel character was maintained when the model was blown by bellows through a rubber tube.

The additional resonator (No. 7) (1024) was added to a resonator having an orifice 23×13 (oval) and resonance 574. The front orifice of the additional resonator was enlarged to 15 mm., giving resonance 861/342. The capacity of the back resonator was reduced, giving 861/383, the front cavity was enlarged giving 608/304 to 322, the back cavity was reduced giving 608/362, the lateral orifice was replaced by a front orifice. The front resonator when separated from the back and its central orifice closed gave 541. This model (No. 9) produced a good u (fig. 12).*

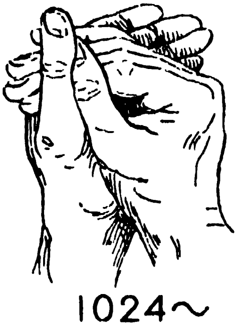


Fig. 9.

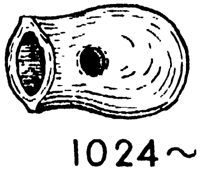


Fig. 10.

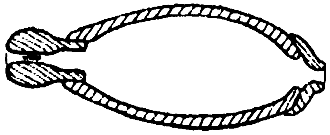


Fig. 11.

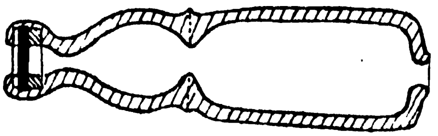


Fig. 12.

Fig. 9.—Clasped hands resonator 1024.
Fig. 10.—Model No. 7—Resonator 1024 in plasticine.
Fig. 11.—Model No. 8—Single resonator (50×100 mm.) with larynx.
Fig. 12.—Model No. 9—u (who) 608/362 (resonators in series).

A similar result was obtained with the additional resonator held in front of the ᴜ model—as in the case of the mouth-formed ᴜ.

From these experiments it appears that:—

(1) u could not be produced by a larynx and a single resonator.

(2) That it can be produced by a larynx and two resonators in series, with the larynx attached either to the back resonator or between the resonators.

---

* It was found that in this model the upper resonance was produced by the back cavity and *vice versa*.

(3) That in blowing a single resonator with the mouth the oral cavity may form a back resonator so as to give a vowel.

(4) That the upper resonance may be formed by the back cavity and the lower resonance by the front.

The front orifice of Model No. 8 was enlarged to 20 mm. and the model blown (with larynx), it then gave a recognisable ei (as in hay).

Experiment was made with this resonator, and it was found that if the orifice was (1) closed, (2) opened, (3) closed, (4) partially opened by hand the model said " baby." A plasticine hood (Model No. 10) was therefore made to form a partial closure or hood (see fig. 13) over the 2-cm. orifice. The model gave a fairly good resonant i free from sibilance, and best heard as mi mi when the opening was intermittently covered and uncovered by the fingers during blowing.

To give a freer passage for the larynx sound, a flat funnel-shaped front resonator was substituted for the hood and the back orifice of the funnel reduced to 2 mm. diameter (Model 11). This produced a greatly improved i with more resonance, which was further improved by shortening the funnel by 1 to 2 mm. giving resonance 2579/322 (see fig. 14).

It was found that further shortening the funnel (so as to raise the upper resonance) reduced the i character, and that adding 10 to 20 mm. altered the vowel to between æ (hat) and ə (sofa); also that a circular-sectioned orifice of the same cross-section as the flat slit and giving the same resonances produced an equally good i.

It appeared therefore that to produce a good i the front resonator must have a relatively small inner orifice, so as to allow the expansion of the air current in the front resonator, but that the sectional shape of this orifice is immaterial.

æ (hat) Model No. 12 (fig. 15), optimum 1932/724. A funnel-shaped model was made about 18 cm. long of cross-section 70 × 40 mm. in front, and 50 × 30 mm. at back, with larynx attached, the mouth of the model being reduced to 40 × 11 mm. This gave 1932/541.

The model was tuned so as to raise the lower resonance without altering the upper resonance, viz.: by enlarging the mouth and the front resonator, and at 1932/645 the model gave a good æ (hat)—heard as uæ uæ when the mouth of the model was closed and released by hand.

Rectangular u model, No. 13 (fig. 16).

With a view to building a model in wood, a rectangular model was made in plasticine with front cavity 70 mm. long, 40 mm. wide, 40 deep, with interior orifice 11 mm., exterior orifice 16 mm., giving resonance of 541 with the interior orifice closed (*cf.* Model No. 9). A rear cavity 73 mm. long × 20 deep × 40 wide with rear orifice 15 mm. (for attachment of larynx) was made, giving 812 with its front (central) orifice closed.

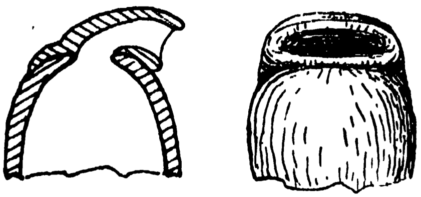


FIG. 13.

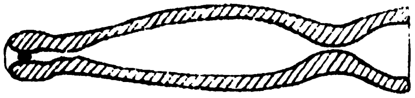


FIG. 14.

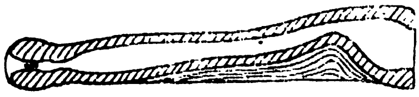


FIG. 15.

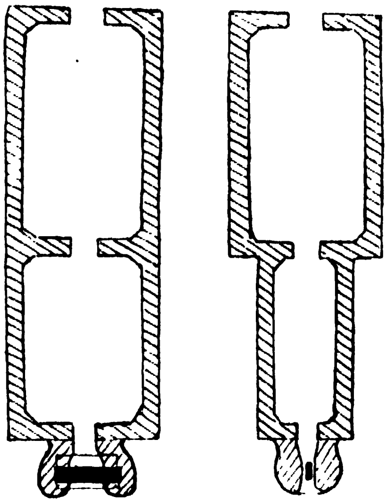


FIG. 16.

FIG. 13.—Model No. 10—Hood on model No. 8 to give i (eat). Section and elevation.
FIG. 14.—Model No. 11—i (eat) model 2579/322.
FIG. 15.—Model No. 12—æ (hat) model 1932/645.
FIG. 16.—Model No. 13—Rectangular u (who) model.

This gave a good u, which was further improved by adding 18 mm. to the length of the front cavity, giving $\left\{\begin{array}{c}645/383\\ \text{Back/Front.}\end{array}\right.$

It is clear from this model that in this case, as in that of No. 9, the higher resonance is given by the back resonator, *i.e.* that the two resonator scales have crossed. A similar model, 13A, was made in wood, but with all linear dimensions doubled; this gave a sound between ʊ and u, but of vocal quality not appreciably different from the plasticine models.

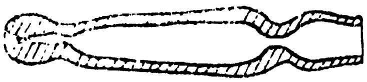


FIG. 17.—Model No. 14—ei (hay) model 2048/483.

ei (hay). Model No. 14 (fig. 17). Optimum resonances 2048 to 2298/430 to 512.

A plasticine back resonator was made 90 mm. × 40 mm. with central orifice 8 mm., giving (with larynx) 512. It was attached to a funnel-shaped front resonator with corresponding central orifice (8 mm.) and front orifice (mouth) about 20 mm. diameter, giving 2298.

The attachment of the front resonator lowered the resonance of the back resonator by five semitones, and that of the front resonator by two semitones. This was corrected by enlarging the central orifice and slightly lengthening the front resonator. The model then gave 2048/483, heard by tapping the resonators, and when blown gave a good ei (as in hay, first part of diphthong) at the first trial.

It was now clear that recognisable reproductions of vowel sounds could be made by passing a larynx note through two resonators in series, such that when joined they resonate respectively to the two characteristic resonances heard in the voice when the vowel is breathed.

ʌ (up). Model No. 15 (fig. 18). Optimum 1448 to 1625/812.

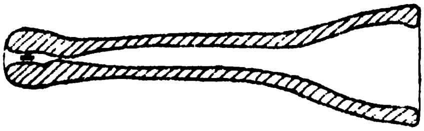


FIG 18.

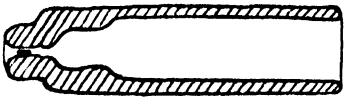


FIG. 18A.

FIG. 18.—Model No. 15—ʌ (up) model 1534/812.
FIG. 18A.—Model No. 15A—ʌ (up) model (cylindrical) 1625/812.

A back resonator with larynx was tuned to 966 (allowing for three semitones drop) with central orifice 25 mm. A front resonator was tuned to 1625 and attached. The front resonator (which was lowered in pitch by attachment) was raised by opening the mouth and slightly reducing its length. The back resonator was found to be one semitone too high. This was corrected by slightly expanding the cavity. The model then gave 1534/812, and when blown gave a good ʌ.

ɔ (all). Model No. 16. Optimum 812 to 966/512 to 608, say 812/512.

A back resonator 60×50×70 mm. with central orifice 15 mm. giving (with larynx) 541, was attached to a front resonator of 100×55×55 mm., with front orifice 35 mm. and rear (central) orifice 15 mm., giving 912. When joined the model gave 861/430, *i.e.* the front resonator was lowered four semitones and the back resonator was lowered one semitone. The model was tuned to 812/512, by reducing the back resonator and expanding the front, and when blown gave a good ɔ (all).

ou (first part of diphthong) (no). Model No. 16 (fig. 19).

Optimum { 683-912/383-483.
Back/Front.

A back resonator (with larynx) with central orifice 20 × 25 mm. was tuned to 812. A front resonator about 70 × 55 mm. cross-section and 125 mm. long, with orifice 24 mm. and back (central) orifice 20 × 25 mm., was tuned to 645. When joined the back resonator was unchanged in pitch (812), the front resonator fell 10 semitones to 362.

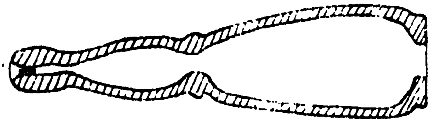


Fig. 19.

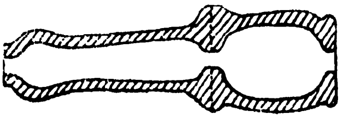


Fig. 19a.

Fig. 19.—Model No. 16—ou (no) model 406/861.
Fig. 19a.—Model No. 17—ou (no) model (resonances reversed) 912/456.

The model was tuned to 812/383 by reducing the front resonator, and gave a good ou (first part of diphthong in no).

ou 456/912 Model No. 17.

A second ou model (No. 17) was made of higher pitch and with the resonances reversed, *i.e.* with a back resonance of 456 and a front resonance of 912. This when blown also gave ou, but tending towards ɔ (as was to be expected, owing to the higher resonances employed). From this experiment it appears that the same vowel sound may be produced by reversing the order of the resonances.

ɒ (not). Model No. 18. Optimum 1084/655.

A back resonator with larynx was tuned to 683 with central orifice about 17 mm.

A front resonator with back orifice 17 mm. was tuned to 1149, with a front orifice of about 30 mm. When joined the lower resonance fell four semitones to 541, the upper resonance fell one semitone to 1084. The lower resonance was raised (by reducing its capacity) to 645.

The model when blown gave ɒ, but rather inclining towards ɑ; the mouth was reduced to 28 mm., lowering both resonances one semitone to 1024/608. This somewhat improved the vowel character.

e (men). Model No. 19. Optimum 2048/574.

A back resonator (with larynx) was tuned to 683 (to allow three semitones drop) with central orifice 25 mm. An open-mouthed front resonator with central orifice 25 mm. was tuned to 2048 with funnel-shaped front orifice about 50 mm. diameter. When joined the lower resonator fell three semitones to 574, the upper fell two semitones to 1824. The front resonator was removed and retuned to 2298, by reducing the length of the funnel, to allow for two semitones drop. On rejoining the resonators the model gave 2048/574 and sounded a good e.

I (it). Model No. 20. Optimum 2169/362.

A back resonator (with larynx) giving 383 (to allow one semitone drop), with central orifice 11 mm. (oval), was attached to a funnel-shaped resonator with central orifice 11 mm. and front orifice 47 mm., giving 2298 to allow one semitone drop. When joined the model gave 2169/362 and a good I.

ɐ (earth) 2nd edition. Model No. 21. Optimum 1534/512.

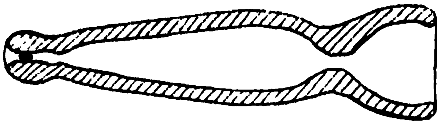


FIG. 20.

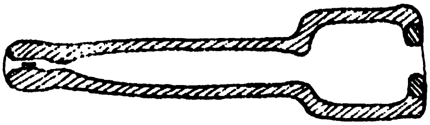


FIG. 21.

FIG. 20.—Model No. 21—ɐ (earth) (2nd edition), 1534/512.
FIG. 21.—Model No. 22—ɑ (calm) (2nd edition), 1366/724.

In view of the inaccurate tuning of the original mouth shaped ɐ model (No. 2) a new model was made, with resonances before joining 1448 and 683, allowing one semitone and five semitones drop respectively, central orifice 15 to 18 mm., front orifice 45 mm.

On joining, the front resonance did not vary, while the back resonance dropped five semitones to 512. The front resonance was raised one semitone by reducing the length of the front resonator.

The model then gave 1534/512 and sounded a good ɐ (earth).*

ɑ (calm) 2nd edition. Model No. 22. Optimum 1217/724.

A back resonator 912 (allowing four semitones drop) with central orifice 25 mm. was joined to a front resonator 1084 with front orifice 23 mm. The back resonance dropped seven semitones. The back resonator was shortened and reduced, and the central orifice slightly enlarged to give 1084.

The front resonator was tuned up to 1217 after joining—by enlarging the front orifice and slightly reducing the capacity. This gave a good ɑ of much better quality than the original model. It may be noted that in the case of "ɑ" the two resonators *before* joining had approximately the *same* resonances.

Effect of Joining Resonators in Series.

The effect of joining two resonators, the one with larynx attached at one end and the other end open, and the other with both orifices open, is in general to lower the resonance of each, but in different proportions according to their relative capacities, and to the relative areas of the common central orifice and the front orifice of the front resonator. Thus, when a large front resonator giving 645 with front orifice 24 mm. diameter and back orifice about 22 mm. was joined to a small back resonator, 812, the front resonator dropped 10 semitones, while the back resonator was practically unchanged (Model No. 16). Conversely when a small front resonator, 2298, with large front orifice, 20 mm., was joined to a large back resonator, 512, with small orifice, 8 mm., the front resonator fell two semitones, while the back resonator fell five semitones. (Model No. 14), *i.e.* for a front resonator, the maximum fall occurs when the front resonator is large and the back resonator small, and the central orifice is large compared with the front orifice.

---

* The sound of this model having been mistaken by many observers for ei, the resonances have been retuned (by reducing the front orifice to 35 mm.) to give 1366/512. This ha improved the ɐ character.

The minimum fall occurs when the front resonator is small compared with the back resonator, and the front orifice is large compared with the central orifice.

For a back resonator the maximum fall is produced by a large back resonator and small front resonator, with a large central orifice and small front orifice, while the minimum fall is produced by a small back resonator combined with a large front resonator. In this case the size of the central and front orifices respectively is practically immaterial.

No exact investigation of the reaction between resonators was made, it being sufficient for the purpose in hand to make an approximate allowance and to correct the resonances after joining up.

Vowel Production by a Single Resonating Cavity.

It was noticed that, in the case of Λ at $\underset{\sharp g''''/\sharp g''}{1625/812}$, a good vowel was produced by the use of a single tubular resonator. Experiment was therefore made with a tubular resonator 43 mm. diameter by 21 cm. long (internal), and one end closed except for a small aperture for blowing the model. The tube was progressively shortened, and the following resonances and vowel characters noted :--

Length.	Resonances.	Difference.	Vowel character (breathed).
Cm.		Semitones	
21	1084 ♯c‴ and 383 g′	18	ɐ (earth).
19	1217 ♯d‴ „ 430 a′	18	ɐ
17	1448 ♯f‴ „ 483 b′	19	ɐ
15	1534 g‴ „ 512 c″	19	Between ɐ and æ (hat).
13	1824 ♯a‴ „ 608 ♯d″.... ..	19	æ
11	2048 c⁗ „ 683 f″	19	æ (upper res. faint).
9	2434 ♯d⁗ „ 812 ♯g″. . . .	19	Between æ and ʌ (up) (upper resonance clear).
7	(1366) f‴ „ 966 + b″ ...	—	ʌ clear, upper resonance inaudible, 1366 heard on blowing.
5	(6134) ? g′′′′′ (1149) d‴ 1366 f‴	—	ʌ inclining to ɐ—a high sibilant g heard.
3	(1625) ♯g‴ 1932 b‴	—	ɪ (it). A high sibilance heard, also a lower sibilance at 1625.

It appears therefore that a cylindrical resonator closed at one end gives at least two audible resonances when blown through, and that certain of these, *i.e.* ɐ at 1448/483, æ at 1824/608, fall within the writer's charted vowel resonances, while ʌ at 1366/966 approaches within three semitones of the charted voice resonances at 1366/812. A principal resonance at 1932 gives an appreciable breathed ɪ (it), within two semitones of 2169 in the voice.

To the above extent therefore, it is possible to produce vowel sounds by a single (cylindrical) resonator.

Experiment was also made with an ovoid plasticine resonator, 60 mm. internal diameter, 105 mm. long, with orifice 29 mm. diameter, giving a principal resonance of c″ (512 ∾). When tapped at its equator the model gave only 512. When tapped at or near either pole, or when blown across the orifice it gave c″ 512 and c⁗ 2048.

A slit 10 × 2 mm. was made at the pole opposite to the orifice and the model was blown, through the slit. It then gave ♯c″ 541, together with ♯g‴ 1625 and ♯c⁗ 2169 with a breathed vowel character like ə (sofa)—(*cf.* the charted ranges *541*-683/1366-*1625*).

When blown with a larynx there was a clear vowel sound between e (men) and ə (sofa), the former character being presumably due to the components 541/2169 (*cf.*, 541/2048 charted for e) and the latter to 541/1625.

It is clear therefore that a single Helmholtz resonator with opposite orifices may, when blown *through*, give two or more resonances so as to produce vowel-like sounds.

Early in the investigation (26th April, 1922) it was recognised that it might be possible to substitute electrical resonators for acoustic resonators, so as to produce vowel sounds in a telephone by passing an intermittent current through two resonating circuits and thence to the telephone. It was also realised that, in view of the production of vowels by a larynx between two resonators (see reference to a experiments with Model No. 8) it was possible that vowels might be produced by passing a larynx note through two resonators in parallel. The electrical production of all the English vowels and of many consonants has since been carried out by John Q. Stewart in the Research Laboratory of the Western Electric Co., of New York (Letter to 'Nature,' dated July 8, 1922, No. 2757, Vol. 110, p. 311), by the use of two electrical resonators in parallel, thus confirming the theory of double resonance.

Resonators in Parallel.

Two plasticine resonators were joined in parallel with their rear orifices adjacent, so as to be conveniently attached to a single larynx, and with their front orifices lying in the same plane; they were tuned to 1534 and 812 respectively. There was no pitch reaction between the resonators.

When blown without larynx the model gave a good breathed Λ (up) and a good voiced Λ when a larynx was added (Model No. 23, fig. 22).

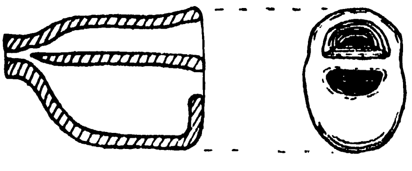


Fig. 22.

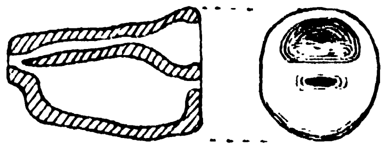


Fig. 23.

Fig. 22.—Model No. 23—Λ (up) by resonators in parallel, 1534/812.
Fig. 23.—Model No. 24—i (eat) by resonators in parallel, 2434/342.

A second model (No. 24, fig. 23) was made with a funnel-shaped resonator giving 2434 (connected by a relatively long passage to the larynx orifice) and a small orificed resonator giving 342. This gave a good breathed i (eat) without larynx and a good, though sibilant, voiced i with larynx.

It was noticed that the 2434 resonator gave a recognisable i when blown *by itself* without larynx, but that when blown by itself with larynx the i character was partly lost, the vowel sound being very poor and more like e (men).

It is clear therefore that vowel sounds may be produced by passing a larynx note through pairs of resonators placed either in series or in parallel.

The identification of the vowel sounds produced by blowing the various models (except in the case of Models 4–10, which were modified after trial) rests on the testimony of more than one observer.

On the Steady Motion of a Cylinder through Infinite Viscous Fluid.

By ARTHUR BERRY, M.A., Fellow of King's College, Cambridge, and LORNA M. SWAIN, Fellow of Newnham College, Cambridge.

(Communicated by H. W. Richmond, F.R.S. Received October 9, 1922.)

§ 1.

There are a good many known solutions of problems of the three-dimensional motion of an infinite viscous fluid disturbed by a moving solid. The simplest of the corresponding two-dimensional problems, that of a circular cylinder moving with uniform velocity, was shown by Stokes[*] to be impossible, when the equations of motion are simplified by the omission of the so-called "inertia terms"; and a general physical argument, given by him to explain the essential difference between the cases of two- and three-dimensions, suggests that the problem is insoluble for a cylinder of any ordinary form, if the "inertia terms" are neglected.

Such solutions as have been given are avowedly incomplete. A recent solution by Lamb[†] of the problem of the circular cylinder, based on Oseen's method,[‡] retains some of the "inertia terms," but omits others. A paper by Wilton[§] involves, in the case of the elliptic cylinder, infinite vorticity and indeterminate velocities at the ends of the axes.

In the present state of knowledge it seems worth while to give another imperfect solution of the problem, when the moving body is an elliptic cylinder, including as special cases a circular cylinder and a plane lamina, either along or perpendicular to the direction of motion.

The inertia terms are neglected; the condition of no slipping at the surface of the cylinder is satisfied exactly; the velocity at infinity in one direction is infinite. As, however, this velocity is only logarithmically infinite, *i.e.*, tends to infinity as the logarithm of the distance from the cylinder, it only increases slowly with this distance, and the solution may be expected to give a fairly good approximation to the motion at moderate distances from the cylinder.

The problem is first treated (§ 2) as a limiting case of the motion of an ellipsoid through infinite viscous fluid, and it is interesting to note that this solution, which fulfils all the boundary conditions, in particular making the velocity zero at infinity, gives, in the limiting case of the elliptic cylinder, a solution which involves the velocity being logarithmically infinite in the direction of flow. The solution would thus appear to be unique, subject to this condition, and we should note that we obtain a definite value for the resistance. The circular cylinder is then treated as a limiting case of the elliptic cylinder.

---

[*] G. G. Stokes, 'Mathematical and Physical Papers,' vol. 3, p. 65.

[†] 'Phil. Mag.,' vol. 21, p. 120 (1911); or Lamb, 'Hydrodynamics,' 4th edit., p. 604.

[‡] 'Arkiv för Matematik, Astronomi og Fysik,' vol. 6, No. 29 (1910).

[§] 'Phil. Mag.,' vol. 30, p. 767 (1915).

In the second place (§ 3) the equations of motion are solved directly and the constants in the solution found by making the velocity vanish on the cylinder and that in the direction of flow be logarithmically infinite at infinity, while the other component is finite at infinity.

Finally, the stream lines, the curves showing the variation of velocity along the stream lines and the curves of constant velocity are drawn for the three limiting cases:—

I. A plane lamina along the stream.
II. „ „ across „
III. A circular cylinder.

The first two are obtained at once by making the minor axis of the elliptic cylinder tend to zero, while the third is obtained by a more difficult limit process.

§ 2.

The solution of the problem of the steady translation of an ellipsoid through viscous liquid is given in Lamb's 'Hydrodynamics' (4th edit.), p. 593. We quote the result with the notation used there. For the ellipsoid $x^2/a^2+y^2/b^2+z^2/c^2=1$, the component velocities (u, v, w), relative to the ellipsoid, are found to be

$$u = \mathrm{A}\,\partial^2\Omega/\partial x^2+\mathrm{B}\,(x\,\partial\chi/\partial x-\chi)+\mathrm{U},$$

$$v = \mathrm{A}\,\partial^2\Omega/\partial x\,\partial y+\mathrm{B}x\,\partial\chi/\partial y.$$

$$w = \mathrm{A}\,\partial^2\Omega/\partial x\,\partial z+\mathrm{B}x\,\partial\chi/\partial z,$$

where

$$\chi = abc\int_\lambda^\infty \frac{d\lambda}{\Delta},\quad \frac{x^2}{a^2+\lambda}+\frac{y^2}{b^2+\lambda}+\frac{z^2}{c^2+\lambda}=1,\quad \Delta^2=(a^2+\lambda)(b^2+\lambda)(c^2+\lambda),$$

$$\Omega = \pi abc\int_\lambda^\infty\left(\frac{x^2}{a^2+\lambda}+\frac{y^2}{b^2+\lambda}+\frac{z^2}{c^2+\lambda}-1\right)\frac{d\lambda}{\Delta},\ \alpha = abc\int_\lambda^\infty\frac{d\lambda}{\Delta(a^2+\lambda)},$$

$$\pi\mathrm{A} = -\tfrac{1}{2}\mathrm{B}a^2,\quad \mathrm{B}=\mathrm{U}/(\chi_0+\alpha_0 a^2),$$

$$\chi_0 = abc\int_0^\infty\frac{d\lambda}{\Delta},\quad \alpha_0 = abc\int_0^\infty\frac{d\lambda}{\Delta(a^2+\lambda)}.$$

To obtain the solution for the elliptic cylinder, we replace the infinite upper limit of χ by an arbitrary constant ξ_1, U by an arbitrary constant C and then make $c \to \infty$. Omit the "w" equation. Introduce the usual variables ξ, η for the elliptic cylinder by writing

$$x = k \cosh \xi \cos \eta, \qquad y = k \sinh \xi \sin \eta\,;$$

and let the ellipse itself be given by $\xi = \xi_0 = \tanh^{-1} b/a$, where $a > b$. Then

$$\chi = ab \int_{\xi}^{\xi_1} \frac{d\lambda}{\sqrt{\{(a^2+\lambda)(b^2+\lambda)\}}} = 2ab \int_{\xi}^{\xi_1} d\xi = 2ab(\xi_1 - \xi),$$

since
$$\lambda = a^2 \sinh^2 \xi - b^2 \cosh^2 \xi,$$

and
$$\alpha = ab \int_{\lambda}^{\infty} \frac{d\lambda}{(a^2+\lambda)^{3/2}(b^2+\lambda)^{1/2}} = \frac{ab}{k^2} \int_{\xi}^{\infty} \frac{2d\xi}{\cosh^2 \xi} = \frac{2ab}{k^2}(1 - \tanh \xi).$$

From these, the various expressions for the velocities can be evaluated and we finally obtain

$$u = \mathrm{A}\frac{4\pi ab}{k^2}\left\{1 - \tanh \xi - \frac{\tanh \xi \cos^2 \eta}{\cosh^2 \xi \sin^2 \eta + \sinh^2 \xi \cos^2 \eta}\right\}$$
$$+ \mathrm{B}\left\{2ab\xi - 2ab\xi_1 - \frac{2ab \cosh \xi \sinh \xi \cos^2 \eta}{\cosh^2 \xi \sin^2 \eta + \sinh^2 \xi \cos^2 \eta}\right\} + \mathrm{C},$$

$$v = -\mathrm{A}\frac{4\pi ab}{k^2} \frac{\sin \eta \cos \eta}{\cosh^2 \xi \sin^2 \eta + \sinh^2 \xi \cos^2 \eta} - \mathrm{B}\,2ab \frac{\cosh^2 \xi \cos \eta \sin \eta}{\cosh^2 \xi \sin^2 \eta + \sinh^2 \xi \cos^2 \eta}.$$

Now $u = v = 0$ on the ellipse $\xi = \xi_0 = \tanh^{-1} b/a$,

therefore

$$\mathrm{A}\frac{4\pi ab}{k^2}\left(1 - \frac{b}{a} - \frac{b}{a} k^2 \frac{\cos^2 \eta}{a^2 \sin^2 \eta + b^2 \cos^2 \eta}\right) + \mathrm{B}\left(2ab(\xi_0 - \xi_1) - \frac{2a^2b^2 \cos^2 \eta}{a^2 \sin^2 \eta + b^2 \cos^2 \eta}\right)$$
$$+ \mathrm{C} = 0,$$

and
$$\mathrm{A}\,2\pi/k^2 + \mathrm{B}\,a^2/k^2 = 0.$$

These are both satisfied by

$$\mathrm{A} = -\mathrm{B}a^2/2\pi,$$

and
$$-2\mathrm{B}(a^3b/k^2)(1 - b/a) + \mathrm{B}\,2ab(\xi_0 - \xi_1) + \mathrm{C} = 0.$$

Hence

$$\frac{u}{\mathrm{B}} = -\frac{2a^2b^2}{k^2} + 2ab(\xi - \tanh^{-1} b/a) + \frac{2ab}{k^2} \frac{\sinh \xi \cosh \xi (a^2 - k^2 \cos^2 \eta)}{\cosh^2 \xi \sin^2 \eta + \sinh^2 \xi \cos^2 \eta}, \qquad (1)$$

$$\frac{v}{\mathrm{B}} = \frac{2ab}{k^2} \frac{(a^2 - k^2 \cosh^2 \xi) \sin \eta \cos \eta}{\cosh^2 \xi \sin^2 \eta + \sinh^2 \xi \cos^2 \eta}. \qquad (2)$$

It will be noticed that $u \to \infty$ as $\log x$, but v remains finite at infinity.

Now
$$\psi = \int (v\,dx - u\,dy),$$

therefore
$$\psi/\mathrm{B} = (2aby/k^2)\left[ab - b^2 \coth \xi + k^2(\xi_0 - \xi)\right]. \qquad (3)$$

To find the resistance per unit length, we require to evaluate

$$\int (lp_{xx} + mp_{xy})\,ds,$$

where l, m are the direction cosines of the normal, and the integral is taken round the ellipse.

The resistance, $\quad R = \int ((\cos\eta/a)\,p_{xx} + (\sin\eta/b)\,p_{xy})\,ab\,d\eta. \qquad (4)$

Now $\quad p_{xy} = \mu\,(\partial u/\partial y + \partial v/\partial x),$

$$p_{xx} = -p + 2\mu\,\partial u/\partial x, \qquad \text{where} \quad p = (\mu/ab)\,\partial\chi/\partial x.$$

Thus p_{xy}, p_{xx} can be evaluated and, on substitution in the integral, we find $R = 4\pi\mu\,(2Bab)$.

The particular case of a plane lamina along the stream is obtained if we replace Bb by a new arbitrary constant and then make $b = 0$, $k = a$; omitting a constant factor, we find

$$\psi = -y\xi,$$

$$u = \xi + \frac{\sinh\xi\cosh\xi\sin^2\eta}{\cosh^2\xi\sin^2\eta + \sinh^2\xi\cos^2\eta},$$

$$v = -\frac{\sinh^2\xi\sin\eta\cos\eta}{\cosh^2\xi\sin^2\eta + \sinh^2\xi\cos^2\eta}.$$

The solution for an elliptic cylinder with its minor axis along the stream is obtained in the same way. It will be sufficient to give the results:—

$$\frac{u}{2Bab} = \frac{ab}{k^2} - \frac{a^2\cos^2\eta + b^2\sin^2\eta}{k^2(\cosh^2\xi\sin^2\eta + \sinh^2\xi\cos^2\eta)}\sinh\xi\cosh\xi + \xi - \tanh^{-1}a/b, \qquad (5)$$

$$\frac{v}{2Bab} = -\frac{b^2\sinh^2\xi - a^2\cosh^2\xi}{k^2(\cosh^2\xi\sin^2\eta + \sinh^2\xi\cos^2\eta)}\sin\eta\cos\eta, \qquad (6)$$

$$\frac{\psi}{2Bab} = y\left\{\frac{b^2\tanh\xi}{k^2} - \frac{ab}{k^2} + \tanh^{-1}\frac{a}{b} - \xi\right\}, \qquad (7)$$

where $b > a$ and $x = k\sinh\xi\sin\eta$, $y = k\cosh\xi\cos\eta$,

$$R/2Bab = 4\pi\mu. \qquad (8)$$

The case of a plane lamina across the stream is obtained if we replace Ba by a new arbitrary constant and then make $a = 0$, $b = k$; omitting a constant multiplier, we then have

$$u = \xi - \frac{\sin^2\eta\sinh\xi\cosh\xi}{\cosh^2\xi\sin^2\eta + \sinh^2\xi\cos^2\eta}, \qquad (9)$$

$$v = -\frac{\sinh^2\xi\sin\eta\cos\eta}{\cosh^2\xi\sin^2\eta + \sinh^2\xi\cos^2\eta}, \qquad (10)$$

$$\psi = y(\tanh\xi - \xi). \qquad (11)$$

We now proceed to the case of the circular cylinder. This can be deduced from the solution for the elliptic cylinder.

Let us take the first solution and change the variables by writing $\xi+\xi_0$ in place of ξ; the ellipse itself is then given by $\xi = 0$. We thus have

$$u = -\frac{2a^2b^2}{k^2}+2ab\xi+\frac{2ab}{k^2}\,\frac{\sinh\overline{\xi+\xi_0}\cosh\overline{\xi+\xi_0}\,(a^2-k^2\cos^2\eta)}{\cosh^2\overline{\xi+\xi_0}\sin^2\eta+\sinh^2\overline{\xi+\xi_0}\cos^2\eta},$$

$$v = \frac{2ab}{k^2}\,\frac{(a^2-k^2\cosh^2\overline{\xi+\xi_0})\sin\eta\cos\eta}{\cosh^2\overline{\xi+\xi_0}\sin^2\eta+\sinh^2\overline{\xi+\xi_0}\cos^2\eta},$$

$$\psi = \frac{2aby}{k^2}\,(ab-b^2\coth\overline{\xi+\xi_0}-k^2\xi).$$

We pass to the circular cylinder by making $\xi_0\to\infty$, $k\to 0$, so that $k\cosh\xi_0\to a$, $k\sinh\xi_0\to a$, where a is the radius of the cylinder.

Then $\qquad x+iy = z = ae^{\zeta} \qquad$ or $\qquad \zeta = \log(z/a)$,

therefore, $\qquad \eta = \tan^{-1}y/x, \qquad \xi = \log(r/a)$.

Now

$$u = 2ab\xi$$
$$+\frac{2ab}{k^2}\,\frac{(a\sin^2\eta\cosh\overline{\xi+\xi_0}-b\cos^2\eta\sinh\overline{\xi+\xi_0})\,(a\sinh\overline{\xi+\xi_0}-b\cosh\overline{\xi+\xi_0})}{\cosh^2\overline{\xi+\xi_0}\sin^2\eta+\sinh^2\overline{\xi+\xi_0}\cos^2\eta}.$$

Hence the coefficient of $\sin^2\eta$ in the numerator is

$$(a^2b/2)\,\{\overline{a-b}\,e^{2\overline{\xi+\xi_0}}-(a+b)\,e^{-2\overline{\zeta+\xi_0}}-2b\}$$
$$= (a^2b/2)\,\{k^2e^{2\overline{\xi+\xi_0}}/(a+b)-(a+b)\,e^{-2\xi}/e^{2\xi_0}-2b\}$$

and becomes in the limit $(a^3/2)\,\{4a^2(r/a)^2/2a-2a\} = a^4(r^2/a^2-1)$.

In a similar way, we see that the coefficient of $\cos^2\eta$ in the numerator becomes $a^4(r^2/a^2-1)$.

Thus

$$u = (a^4/r^2)(r^2/a^2-1)\,((y^2-x^2)/r^2)+2a^2\log r/a$$
$$= a^2\{(a^2/r^4-1/r^2)(x^2-y^2)+2\log r/a\}. \tag{12}$$

Similarly,

$$v = \frac{2ab\cos\eta\sin\eta\,\{a^2-(\cosh\xi\,.\,k\cosh\xi_0+\sinh\xi\,.\,k\sinh\xi_0)^2\}}{k^2(\cosh^2\overline{\xi+\xi_0}\sin^2\eta+\sinh^2\overline{\xi+\xi_0}\cos^2\eta)}$$
$$= (2a^4/r^4)\,xy\,(1-r^2/a^2) = a^2\,.\,2xy\,(a^2/r^4-1/r^2). \tag{13}$$

In a similar way, we can prove that

$$\psi = a^2\{y(1-a^2/r^2)-2y\log(r/a)\}. \tag{14}$$

§ 3.

Secondly, it may be interesting to show that the solutions for the elliptic and circular cylinders can be obtained directly from the equations of motion and need not be deduced as particular cases of the ellipsoid.

When the inertia terms are neglected, the stream function ψ for two-dimensional motion satisfies the equation

$$\nabla^4\psi = 0.$$

To obtain a solution, put $\nabla^2\psi = \psi_1$, then

$$\nabla^2\psi_1 = 0.$$

For the circular cylinder, assume $\psi_1 = (\mathrm{A}'r + \mathrm{B}'/r)\sin\theta$[*], where A', B' are constants. Then,

$$\nabla^2\psi = (\mathrm{A}'r + \mathrm{B}'/r)\sin\theta.$$

An integral of this is

$$\psi = (\mathrm{A}r\log r + \mathrm{C}r + \mathrm{D}/r + \mathrm{B}r^3)\sin\theta,$$

where A, B, C, D are all constants.

Hence

$$u' = -\partial\psi/r\partial\theta = -(\cos\theta/r)(\mathrm{A}r\log r + \mathrm{C}r + \mathrm{D}/r + \mathrm{B}r^3),$$

$$v' = \partial\psi/\partial r = \sin\theta(\mathrm{A} + \mathrm{A}\log r + \mathrm{C} - \mathrm{D}/r^2 + 3\mathrm{B}r^2).$$

To have the velocity only logarithmically infinite at infinity, we put $\mathrm{B} = 0$. Then the condition of no slipping on the cylinder $r = a$ leads to the equations

$$\left.\begin{aligned} \mathrm{A}\log a + \mathrm{C} + \mathrm{D}/a^2 &= 0 \\ \mathrm{A} + \mathrm{A}\log a + \mathrm{C} - \mathrm{D}/a^2 &= 0 \end{aligned}\right\},$$

therefore $\qquad \mathrm{D} = \tfrac{1}{2}\mathrm{A}a^2$, and $\mathrm{C} = -\mathrm{A}\log a - \tfrac{1}{2}\mathrm{A}$,

therefore $\qquad \psi = -\tfrac{1}{2}\mathrm{A}\,y\,(1 - a^2/r^2 - 2\log(r/a))$,

and this agrees with formula (14), if we remember that A is an arbitrary constant.

For the elliptic cylinder, if we introduce the usual co-ordinates (ξ, η) defined by

$$x + iy = k\cosh(\xi + i\eta),$$

the differential equation satisfied by the stream function ψ is

$$\left(\frac{\partial^2}{\partial\xi^2} + \frac{\partial^2}{\partial\eta^2}\right)\left[\frac{1}{\cosh^2\xi\sin^2\eta + \sinh^2\xi\cos^2\eta}\left(\frac{\partial^2}{\partial\xi^2} + \frac{\partial^2}{\partial\eta^2}\right)\right]\psi = 0.$$

---

[*] We assume this form, as it will then be possible to make the velocity parallel to the axis of x only logarithmically infinite at infinity, and that parallel to the axis of y finite at infinity.

Put $$\frac{1}{\cosh^2\xi\sin^2\eta+\sinh^2\xi\cos^2\eta}\left(\frac{\partial^2}{\partial\xi^2}+\frac{\partial^2}{\partial\eta^2}\right)\psi = \psi_1,$$

then $$\left(\frac{\partial^2}{\partial\xi^2}+\frac{\partial^2}{\partial\eta^2}\right)\psi_1 = 0.$$

This is satisfied if we take

$$\psi_1 = \text{imaginary part of } [-\tfrac{1}{2}\mathrm{A}'\operatorname{cosech}(\xi+i\eta)],$$

where A′ is a real constant;[*]

$$= \frac{\mathrm{A}'\cosh\xi\sin\eta}{\cosh^2\xi\sin^2\eta+\sinh^2\xi\cos^2\eta}.$$

This form for ψ_1 will allow the boundary conditions on $\xi = \xi_0$ to be satisfied, make the velocity parallel to the axis of x only logarithmically infinite at infinity and that parallel to the axis of y finite.

We then have as an integral

$$\psi = \sin\eta\,(\mathrm{A}\cosh\xi+\mathrm{B}\sinh\xi+\mathrm{C}\xi\sinh\xi),$$

where A, B, C are constants.

The two boundary conditions, $\partial\psi/\partial x = \partial\psi/\partial y = 0$, to be satisfied on $\xi = \xi_0$, lead to the equations

$$(\mathrm{A}\sinh\xi+\mathrm{B}\cosh\xi+\mathrm{C}\sinh\xi+\mathrm{C}\xi\cosh\xi)\sin\eta\,\partial\xi/\partial x$$
$$+(\mathrm{A}\cosh\xi+\mathrm{B}\sinh\xi+\mathrm{C}\xi\sinh\xi)\cos\eta\,\partial\eta/\partial x = 0,$$

$$(\mathrm{A}\sinh\xi+\mathrm{B}\cosh\xi+\mathrm{C}\sinh\xi+\mathrm{C}\xi\cosh\xi)\sin\eta\,\partial\xi/\partial y$$
$$+(\mathrm{A}\cosh\xi+\mathrm{B}\sinh\xi+\mathrm{C}\xi\sinh\xi)\cos\eta\,\partial\eta/\partial y = 0.$$

On reduction, the first gives

$$-\mathrm{A}+\mathrm{C}\sinh^2\xi_0 = 0,$$

and the second

$$\left.\begin{aligned}\mathrm{A}\cosh\xi_0+\mathrm{B}\sinh\xi_0+\mathrm{C}\xi_0\sinh\xi_0 &= 0\\ \mathrm{A}\sinh\xi_0+\mathrm{B}\cosh\xi_0+\mathrm{C}\sinh\xi_0+\mathrm{C}\xi_0\cosh\xi_0 &= 0\end{aligned}\right\}.$$

Of these, only two are independent and the third follows; they give

$$\mathrm{A} = \mathrm{C}\sinh^2\xi_0, \qquad \mathrm{B} = -\mathrm{C}(\xi_0+\sinh\xi_0\cosh\xi_0).$$

Hence

$$\psi = \mathrm{C}\sin\eta\,(\sinh^2\xi_0\cosh\xi-\xi_0\sinh\xi-\sinh\xi_0\cosh\xi_0\sinh\xi+\xi\sinh\xi)$$
$$= -(\mathrm{C}/k^3)\,y\,(ab-b^2\coth\xi+k^2\xi_0-\overline{\xi}).$$

This value agrees with formula (3).

The case of the elliptic cylinder with its minor axis along the stream can be solved in a similar manner by equating ψ_1 to the real part of $[-\tfrac{1}{2}\mathrm{A}'\operatorname{cosech}(\xi+i\eta)]$.

§ 4.

The accompanying figures show (A) the stream lines, (B) the variation of the resultant velocity (q) along the stream lines, and (C) the curves of constant

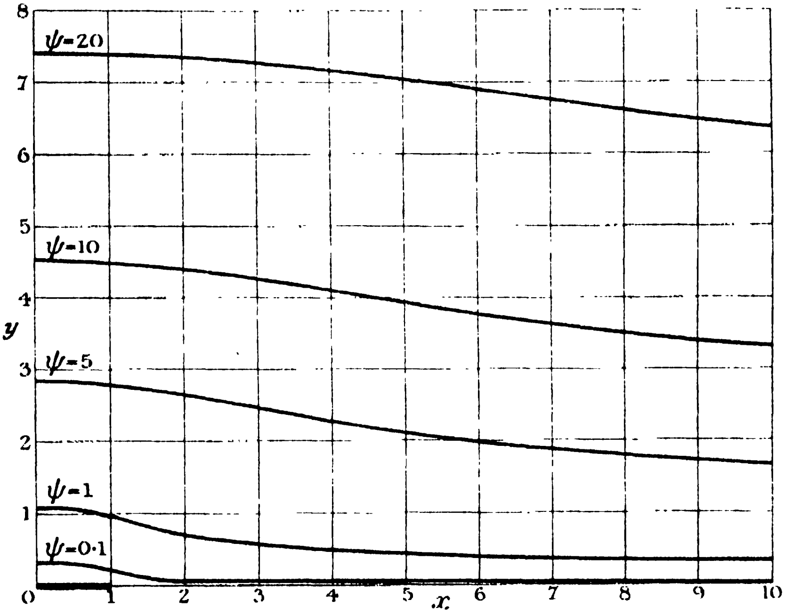


Fig. I (A).—Plane lamina along the stream : stream lines.

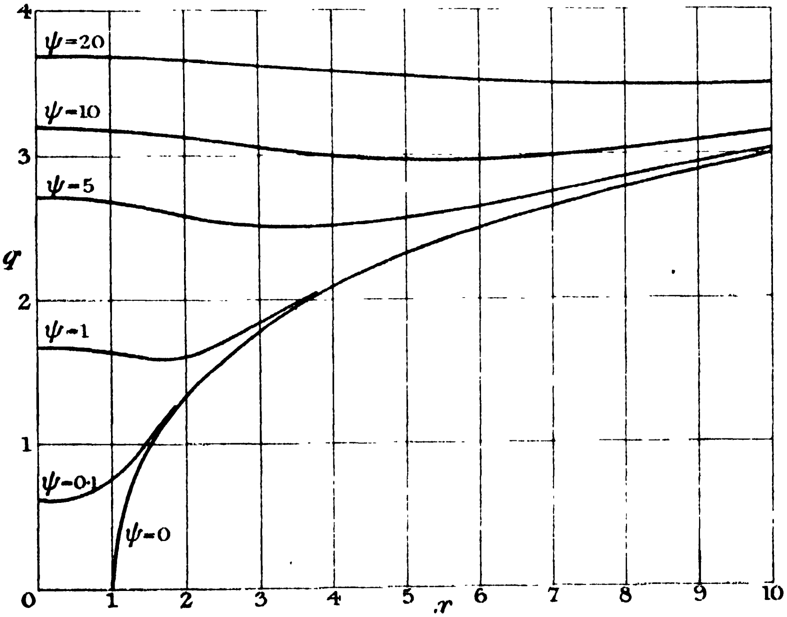


Fig. I (B).—Plane lamina along the stream : variation of velocity along stream lines.

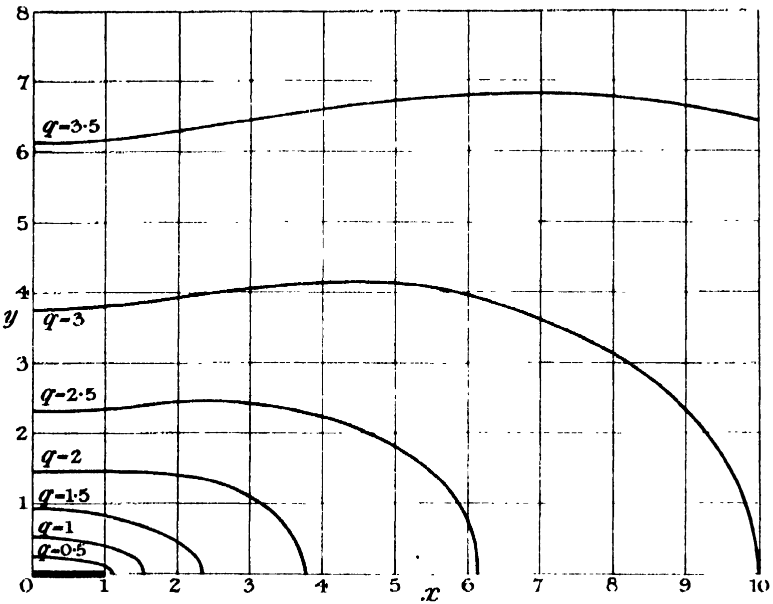


FIG. I (C).—Plane lamina along the stream : curves of constant velocity.

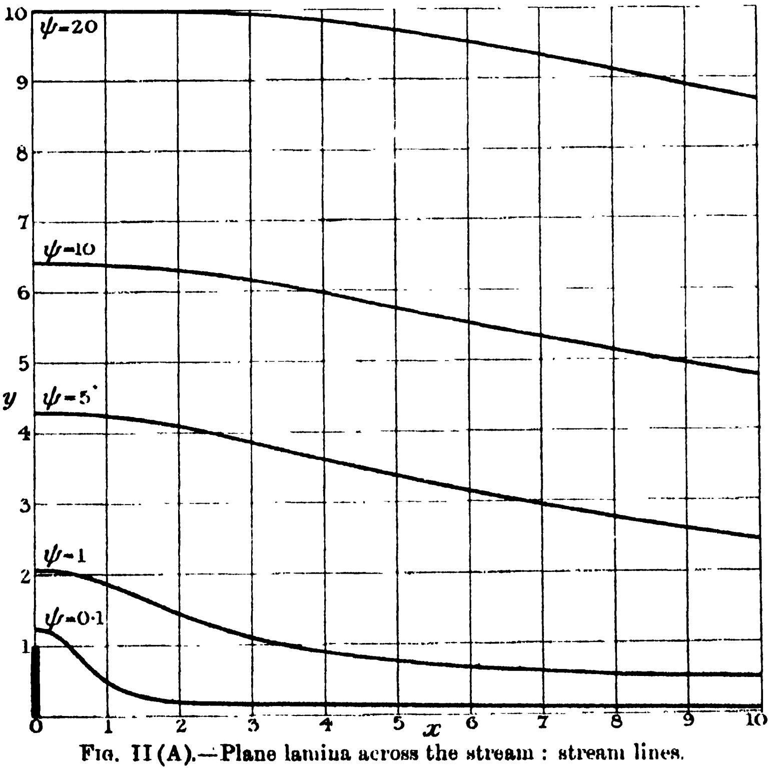


FIG. II (A).—Plane lamina across the stream : stream lines.

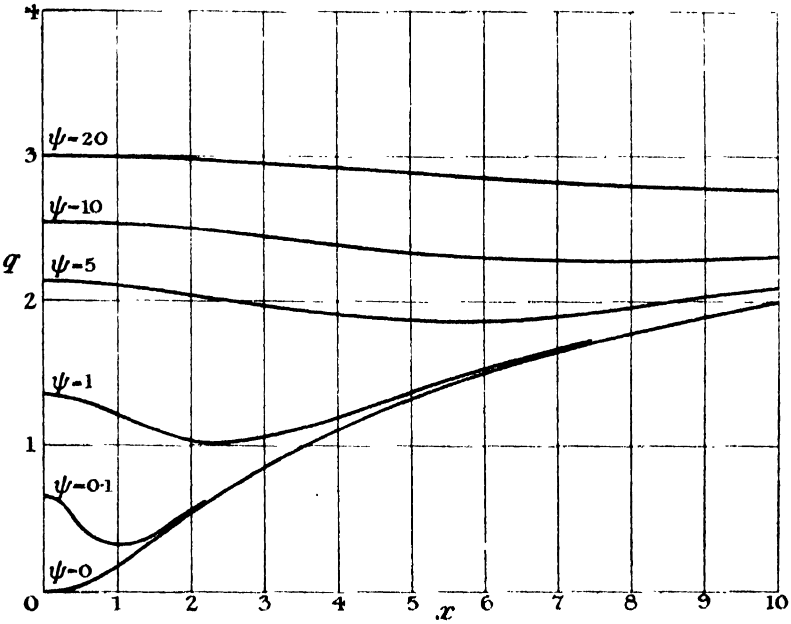


FIG. II (B).—Plane lamina across the stream: variation of velocity along stream lines.

velocity, for each of the cases: (I) a plane lamina along the stream, (II) a plane lamina across the stream, and (III) a circular cylinder. In each case only one quadrant of the field is shown, so that in cases I and II the thick line drawn represents half the lamina.

It will be noticed that the curves are very similar for the three cases and that the stream lines bear a very strong resemblance to those given for a circular cylinder in a stream of finite breadth in a recent paper by Bairstow, Cave, and Lang.* The scale of the figures here given is quite indeterminate, as the solution always contains an arbitrary constant multiplier. For any particular case it could be experimentally determined by making the velocity at some point agree numerically with that given by the theory. We see that along the stream line $\psi = 0$ the velocity steadily increases, but along all the other stream lines it decreases first and then increases, tending rapidly towards the values assumed along the stream line $\psi = 0$. It will be observed that the magnitudes of the velocities tend very slowly to infinity, and that the velocity increases more rapidly as we recede from the obstacle across the stream than with the stream.

---

* "The Two-dimensional Slow Motion of Viscous Fluids," 'Roy. Soc. Proc.,' A, vol. 100, p. 394 (1922). This paper is a continuation of a previous paper "On Two-dimensional Solutions of Poisson's and Laplace's Equations," by L. Bairstow and A. Berry, 'Roy. Soc. Proc.,' A, vol. 95, p. 457 (1919).

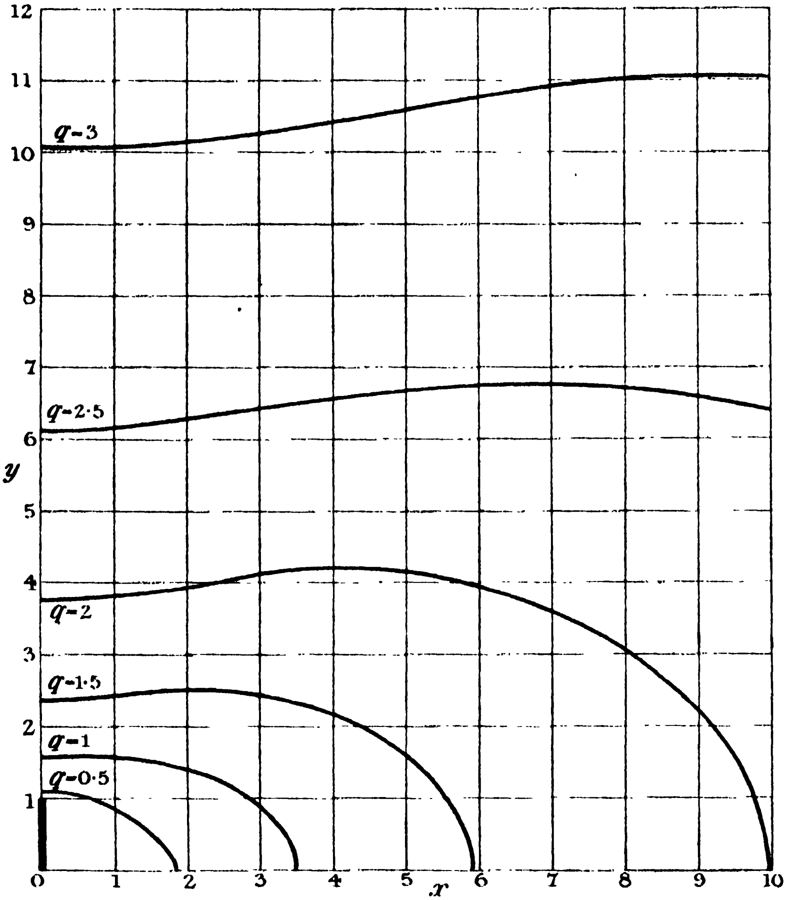


FIG. II (C).—Plane lamina across the stream: curves of constant velocity

To illustrate the slow rate of increase of the velocity, let us choose the radius of the circle, or the semi-length of the lamina, as unit of length and for unit velocity, the velocity at a point on the axis of y at unit distance from the cylinder.

We find for points on the axis of x—

for

$x = 5$, $q = 1·5$ in Case I; $q = 1$ in Case II, $q = 1·1$ in Case III;

$x = 10$, $q = 2·0$ in Case I; $q = 1·5$ in Case II; $q = 1·7$ in Case III.

In illustration of the change of velocity along a stream line, we find:

In Case I, for $\psi = 20$, q decreases slightly between $x = 0$ and $x = 10$,

$\psi = 10$, q is sensibly unchanged,

$\psi = 5$, q increases by 10 per cent.;

In Case II, the corresponding values of q all decrease;
In Case III, for $\psi = 20$, q decreases by about 10 per cent.

$\psi = 10$, q decreases by about 7 per cent.

$\psi = 5$, q decreases by about 10 per cent.

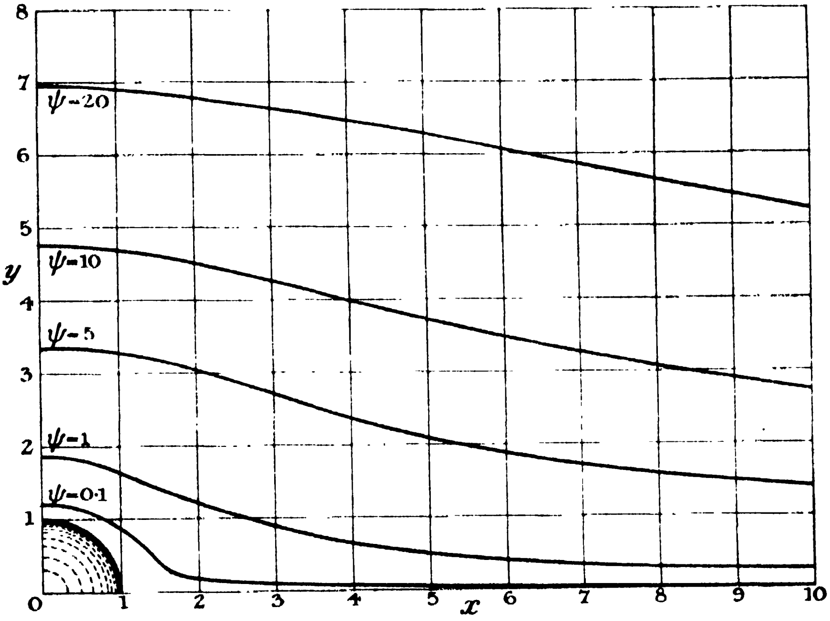


FIG. III (A).—Circular cylinder: stream lines.

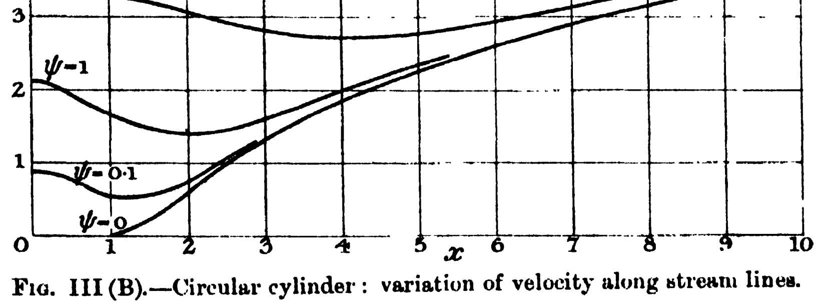


FIG. III (B).—Circular cylinder: variation of velocity along stream lines.

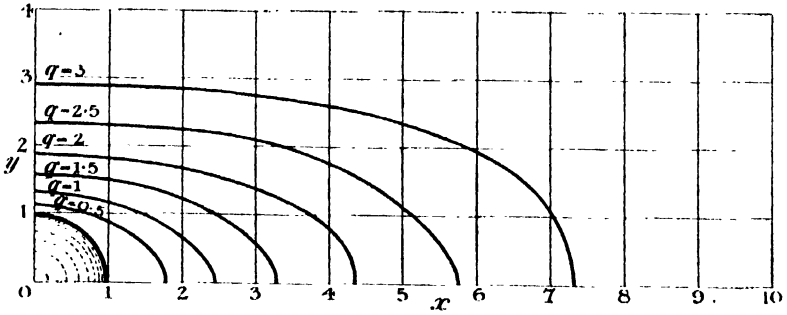


FIG. III (C).—Circular cylinder: curves of constant velocity.

Hence it seems probable that the theory should represent the facts fairly accurately for some considerable distance from the cylinder. The solution will, of course, break down at infinity, as the inertia terms could no longer be neglected, when the velocity is logarithmically infinite.

In conclusion, the authors would like to express their gratitude to Miss E. D. Lang, M.A., to whom they are indebted for much of the arithmetic, and to Dr. G. F. C. Searle, F.R.S., and to Mr. C. F. Sharman, of King's College, Cambridge, who have kindly drawn the curves for them.

OBITUARY NOTICES

OF

FELLOWS DECEASED.

CONTENTS.

PAGE

Sir Alfred Bray Kempe (with portrait) i

Sir William Christie xi

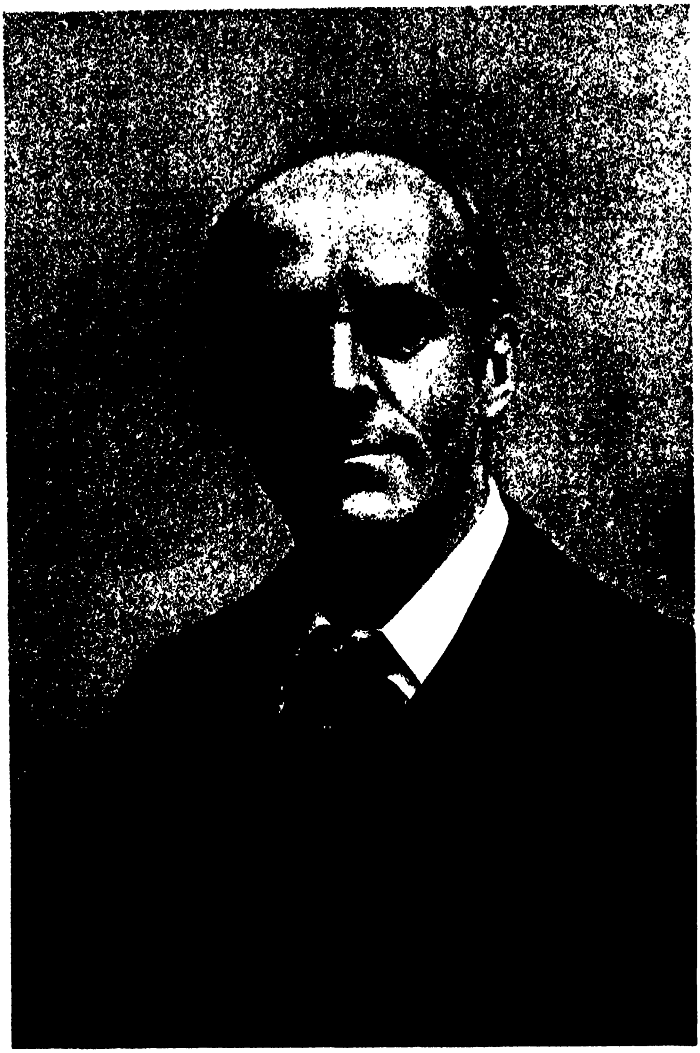
William Gowland (with portrait) xvi

E. H. Grove-Hills xx

G. W. Walker xxii

C. G. Knott xxvii

Jacobus Cornelius Kapteyn xxix



SIR ALFRED BRAY KEMPE, 1849–1922.

AMONG the losses which the Royal Society has recently sustained none has evoked deeper regret than the death of Sir Alfred Bray Kempe, who for twenty-one years, as its Treasurer and one of its Vice-Presidents, took a leading share in the management of its affairs and in the promotion of its prosperity. Some grateful record of his career could not find a more appropriate place than in the pages of the 'Proceedings' of the Society with which he was so long and so closely associated.

The third son of Prebendary John Edward Kempe, Rector of St. James's, Piccadilly, he was born on July 6, 1849. From St. Paul's School, as Camden Exhibitioner, he passed to Trinity College, Cambridge, where, in 1872, he took his degree with special distinction in Mathematics. In the same year he published his first mathematical paper, the title of which—"A general method of solving equations of the *n*th degree by mechanical means"—showed the bent of his mind in scientific enquiry. For some years he continued to publish mathematical essays, but having chosen the Law as his profession, and become a Barrister of the Inner Temple and Western Circuit, he was soon immersed in legal business. To the last, however, he never wholly relinquished his mathematical studies. He used to say of himself that his favourite recreations were Mathematics and Music. He was hardly ever without some problem at which, in such leisure as he could find, he steadily worked. But he refused, as he said, to "empty his note-books into the 'Proceedings' of the Royal Society." He would not be induced to publish his studies until he had really got to the bottom of his enquiry.

In the early part of his legal career, before he became an authority on Ecclesiastical Law, he met with some of the amusing incidents which vary a barrister's experiences, and these he used to tell with great glee. There was one of his stories in which he related how, as a young lawyer, he had been sent to Germany to take evidence for a case in Court, but had his proceedings interrupted by the authorities and was actually arrested on a charge of "usurping the functions of the German Kaiser!" He was soon released, however, on the score of his ignorance of the law—a palliation which he laughingly said was rather hard on a barrister.

Sir Lewis Dibdin, Dean of the Arches, who of all his legal comrades was perhaps his most intimate associate, has been so good as to supply for this record the following recollections of Kempe's life as a barrister.

"I first came in contact with Kempe in 1881 when he was Secretary of the Ecclesiastical Courts Commission which reported in 1883. This important body, included Tait, Benson, Stubbs, Westcott, Freeman, Jeune and many other distinguished members. Kempe impressed them all by his admirable work and knowledge. Chancellorships began to drop in, and he soon became one of the few recognized authorities on Ecclesiastical Law at the Bar. He was one of the Bishop's Counsel in the Lincoln case and also appeared before the Archbishops at the Lambeth Hearing on 'Reservation.' His opinion was sought in a great number of Church cases, and in the much rarer event of proceedings in Court, Kempe was almost certain to be briefed on one side or the other. In 1912 he obtained the 'blue-ribbon' of the Chancellorships—the Chancellorship of London—and as his health was not very strong, ceased to appear as counsel in Court though, I believe, he still advised on 'cases.' In 1913 I asked him and the late Sir Charles Chadwyck Healey, K.C.B., K.C., to act jointly with myself as a committee formed at the request of the Archbishops to ascertain what steps were being taken to secure protection for Church fabrics and to report and make recommendations to their Graces. Kempe contributed a valuable memorandum as to the law of Faculties, which was printed with the Report dated July, 1914. The War probably prevented the Report's recommendations, which were unanimous, from receiving more attention. Kempe's reported judgments in the Consistory Courts over which he presided are not numerous.

"He was an admirable lawyer. His logical mind, coupled with real learning and knowledge of cases, made his opinions clear and sound. It was a pleasure to be associated with him in the consideration of legal questions. While his own arguments were easy to follow, it was equally easy to make him follow those of other people. As an opponent in Court he was not less satisfactory. Always courteous and rigidly fair, he could be relied on to put a winning case convincingly. He was not made for the rough and tumble of contentious advocacy. I think his amiable and refined temperament rather revolted from it, and he was not at his best with a bad case. Probably the clarity of his mind made it difficult for him to argue a rotten point. It would be true to say that he so conducted his own side of a case as always to win the respect of an opponent, while if one had much to do with Kempe, respect inevitably ripened into a warm regard and affection."

The Secretaryship of the Royal Commission on Ecclesiastical Courts lasted for two years. The recognition of the Secretary's growing mastery of Ecclesiastical Law naturally led to the "dropping in" of Chancellorships of Dioceses above referred to. In the course of years Kempe held the Chancellorships of Newcastle, Southwell, St. Albans, Peterborough, Chichester, Chelmsford, and finally that of London which he filled for the last ten years of his life. He also became a Bencher of the Inner Temple. It was in the early part of his career that he found most time to prosecute his researches in the department of mathematics which had specially attracted him. The papers published by him previous to 1879 were the ground on which he was, in that year, proposed for election into the Royal Society. The Certificate wherein his proposers narrated his claims described him as "distinguished for his knowledge of and discoveries in Kinematics." It was signed by a group of the foremost mathematicians of the day, including Cayley, Sylvester, and others. He duly became F.R.S. on June 2, 1881. His friend

Major P. A. MacMahon, F.R.S., well known for contributions to a kindred department of science, has been so good as to furnish for this Obituary the following appreciation of his mathematical work:—

"Kempe's chief contribution to Mathematical Science was his 'Memoir on the Theory of Mathematical Form' ('Phil. Trans. R.S.,' 1886). He considered the subject matter of mathematical thought to consist of a number of differing and non-differing individuals and pluralities, and that the duty of the mathematician is to investigate the characteristics of such matter. Usually the subject matter of thought is accompanied by what he termed 'accidental clothing,' which may be geometrical, algebraical, logical, etc., and his object was to separate it so as to present it in its bare form ready for any raiment that the investigator may find to be appropriate. These bared mathematical forms exist in infinite variety, and any one *may* appertain to subjects of thought which to all appearance have little or nothing in common.

"A classification of *forms*, in the sense used by Kempe, involves the classification of all the matters that may be subjected to mathematical thought and processes. Every subject matter is, in his phrase, reduced to *necessary* subject matter, and he shows wherein consists the infinite variety which the necessary matter exhibits. He studies the nature of the collections of individuals and pluralities, but restricts himself to an exposition of the fundamental principles. After setting forth the definitions and elementary developments, he shows the applications to a variety of cases in such a manner as to vindicate clearly the basic principles of the study. He also describes a simple and uniform method of separating, in any given case, the essential from the non-essential material. He puts in evidence, to this end, a collection of 'units' which may be distinguishable or undistinguishable, and grouped, or not, into pairs, triads, . . . n-ads. He shows that every collection of units has a definite *form* due (i) to the number of its component units; (ii) to the way in which the distinguished and undistinguished units, pairs, triads, etc., are distributed throughout the collection. To quote his own words, units may denote 'material objects, intervals or periods of time, processes of thought, points, lines, statements, relationships, arrangements, algebraical expressions, operators,' etc., etc., and may occupy various positions and be otherwise variously circumstanced. It is thus evident that he took a comprehensive view of the work discussed in the memoir.

"In the course of his masterly development, he brings himself into contact with W. K. Clifford's papers 'On the Types of Compound Statement involving Four Classes' (Proc. Man. Phil. Soc.,' vol. 6, 3rd series), with Grassmann's 'Extensive Algebra' ('Amer. J. of Math.,' vol. 1, pp. 350 *et seq.*), and with Venn's well-known work on 'Symbolic Logic.' In particular, he carries the principles far enough to include primitive and compound algebras, and is able to exhibit the ordinary algebra of quantity as one compounded of two primitive algebras.

"It was the desire to see the subject matter of thought in its absolutely lowest terms that dominated Kempe's mathematical activity during the twenty years (1875–95) in which he published his work. He was President of the London Mathematical Society for two years. When he retired from the Chair in 1895 he recurred again to the subject in his valedictory address. In that discourse he discussed the question—What is Mathematics? He referred to answers to this question that had been given by John Hopkinson, Civil Engineer, and by Venn, Logician, from their special points of view. He quoted De Morgan as saying that 'Space and Time are the only necessary matters of thought, and thus form the subject matter of mathematics,' and Benjamin Peirce as responsible for the statement, 'Mathematics is the science which draws necessary conclusions.' He did not find these replies satisfactory, and doubted whether either was likely to have effect upon the march of mathematical research. He had in mind a definition which might be effective in promoting research by suggesting new paths, new processes, and new classifications and co-ordinations. At the conclusion of his address he is led to the reply upon which he had founded his theory of Mathematical Form nine years before. He did not regard it as being a perfect definition, but as the best that he had been able to devise, and he looked forward to a better one being forthcoming at some future time.

"There is no doubt that his ideas enabled him to visualise a mathematical question, and, indeed, almost any subject of thought, in a novel, interesting and suggestive manner. By means of his graphs of points variously coloured, placed in certain relative positions and connected (or not) by lines, single or multiple, variously distinguished, he was able to form a mental picture of any subject of thought, and to ascertain the nature and extent of the essential differences between different subjects. In this respect he perhaps resembled W. K. Clifford more than any other mathematician ancient or modern. This was recognised by those who had the responsibility of dealing with Clifford's posthumous papers, so that the advice and assistance of Kempe was sought, to the great advantage of science. It should be mentioned that Clifford had visualised, in the Kempe manner, much of the theory of Algebraical Invariants, at that time a comparatively new study, but had met at a certain point with difficulties which at the time of his death he had not succeeded in surmounting. Kempe, however, had gone somewhat deeper than Clifford into the graphical representation of mathematical form, and, moreover, possessed just that knowledge of the theory of Invariants which enabled him within a short time to fill up the lacunæ in Clifford's work. It is safe to say that at that time no one but Kempe could have achieved this. He was the one mathematician qualified for the task.

"Early in his career he was interested in Linkages, a subject which came to the front by reason of the discovery by Peaucellier—an officer in the French Engineers—of a linkage which would draw a straight line. The want of such a mechanism had been felt in engineering practice. Watt, in his steam-engine, had adopted a linkage for guiding his piston, which was an approximate straight-line motion, but the exact solution of the problem had been regarded as an impossibility. Peaucellier's discovery infused new interest into the subject, and other linkages which drew straight lines or arcs of circles of given very large radii were soon forthcoming. Kempe, early in the field, gave a delightful series of lectures in Kensington with the title 'How to Draw a Straight Line,' in which he described Peaucellier's discovery and the subsequent developments. These discourses were published in 'Nature,' with many illustrations, and probably constitute the best existing account, both popular and scientific, of the subject.

"He wrote several other papers, mostly on algebras with particular laws, which all bear the impress of his ability to get down to bed rock in any subject that was occupying his mind. His legal training led him in all cases to lucid and exact statements. His mathematical work, though not large in quantity, was first-rate in quality. What he put forward for publication was his best, and he will be always remembered as a noteworthy contributor to the Philosophy of Mathematics.'

The publication of his mathematical papers established Kempe's reputation as a man-of-science. But the philosophers quickly discovered that he was also an excellent man-of-business. In 1897, he was elected into the Council of the Royal Society, where he soon took a leading place. Accordingly, in the following year, when the Treasurership of the Society became vacant, the Council resolved to propose him for election to this important office. Some of the elder and more conservative Fellows, however, were not prepared to place a comparatively young man in a post which they thought should always be filled by a man of years and of long experience. The general body of the Society supported the President and Council, and Kempe was duly elected Treasurer on St. Andrew's Day, 1898, an office which, combined with that of Vice-President, he worthily held for twenty-one years.

A barrister who in his full professional career would undertake the exacting labours of this Treasurership showed no little courage. Yet from the beginning to the end of his tenure of the office Kempe devoted himself with unstinted zeal not only to his special financial duties, but to the general multiplied business of the Society, as if he had no other vocation in life. It was apt to be forgotten that the time which he gave to the work of the Society was found by him in the midst of all the claims of his profession. The general body of the Fellows thought of him not as the eminent lawyer, but as one of themselves, a notable man-of-science who had undertaken to guide the financial affairs of their Society, and who at each Anniversary gave a brief account of his stewardship. His report on these annual occasions formed one of the most interesting features of the meeting. With great clearness, and often with not a little humour, he would sketch the financial position of the Society, and the state of the funds and investments under his charge. He made himself familiar with the history and purport of the numerous trusts which he had to administer, and he took care to revise and keep up to date the account of each of them given in the Society's 'Record.'

But only his colleagues in the Council could be fully aware of the amount and varied nature of the work which he accomplished for the Society. Besides mastering the business of the Council, he was an active member of many Committees, especially of those which involved expenditure of funds, where his presence as "financial assessor" was of service. His clear common sense, legal knowledge, wide experience of men, and gift of clear exposition, gave to his opinion great influence in the Council, and contributed, in no small measure, to shape the policy and sustain the prosperity of the Royal Society.

It was always interesting to observe with what energy and evident enjoyment he would plunge into a complicated piece of business and gradually reduce the confusion into intelligible order. A remarkable instance of this strong mental grasp was afforded in 1905. At that time the Statutes of the Society, in consequence of successive alterations and additions, stood in such need of revision and consolidation that a special committee was appointed to deal with the matter. The Treasurer found the task to be eminently congenial to him, but to be also more easily accomplished in the quiet of his home than amid the discussions of the committee-room. Taking his copy of the Society's 'Year-Book' he entirely recast the chapters on the Statutes, removing some of them into the class of Standing Orders, and making many alterations, additions and improvements, in the direction of clearness and precision. The volume in which he wrought this transformation, together with the inserted slips and pages of fresh manuscript, has been preserved, and through the kindness of Lady Kempe is now placed among the archives of the Society. A comparison of its contents with the new Section of Statutes which issued from the hands of the Committee shows that the Treasurer's revision was accepted. The great changes which he made can best be appreciated by comparing the Statutes, as altered, with those previously in force.* The President and Council, in reporting to the Fellows the completion of the revision, naturally remained modestly silent as to the Treasurer's share in the work. But his volume, with its crowded corrections and additions in ink and pencil, will remain as an interesting example of the thoroughness of all that he did for the Society.

The boldest step ever taken by the Royal Society was probably its acceptance of the control of the National Physical Laboratory. The urgent need of the establishment of an institution for physical testing and standardisation, in a great industrial and manufacturing country like Britain, had been strongly advocated by Fellows of the Society, more especially by Lord Rayleigh, and eventually the President and Council resolved to take charge of the efforts to create and maintain such a laboratory. A favourable site and a commodious house at Bushey having been obtained from the Crown, the Treasury was induced to place a small grant on the Parliamentary Estimates, and a number of generous subscriptions came from well-wishers throughout the country. It was provided that the control of the Laboratory should be placed in the hands of the President and Council of the Royal Society, and that its income and all other property should be vested in the Society. At the same time, the Society became responsible for any deficit that might occur in the annual income. A spacious laboratory and other buildings were erected, and by March, 1902, various departments were so far advanced that the institution could be formally opened by the Prince of Wales. This initial success of the undertaking was followed by a rapid growth, fresh departments being started year after year. The advance would have been even more marked had the available funds permitted. But the expenses of management sometimes exceeded the income. Hence, though the Laboratory has amply justified the anticipations of its founders, the financial questions arising out of its development were a constant source of solicitude to the President and Council, and more especially to the Treasurer, who from the beginning, as the original Treasurer of the National Physical Laboratory, took a keen interest in its success. Besides watching over the finance of the institution, he did notable service in obtaining subscriptions and otherwise increasing its revenue. The outbreak of the War, in 1914, augmented and complicated the financial difficulties. It was now becoming evident that the task of conducting such a large and costly national institution as the Laboratory had become, lay beyond the province of any scientific society, and ought properly to be undertaken by the State. This transference of control was at last effected in 1918. On April 1 of that year, the Royal Society, having initiated the National Physical Laboratory and fostered its development for sixteen years, had the proud satisfaction of handing over to a Government Department this active and continually growing institution, which had proved itself to be an important addition to the scientific resources of the country. Among those who took a large but unobtrusive share in its development, the name of Alfred Bray Kempe deserves to live in grateful remembrance.

---

* See 'Record of the Royal Society,' 1912, pp. 131-158.

As befitted a barrister, Kempe evidently loved the definiteness, precision, and even the redundancy of legal language: and, as he frequently had to draft a formal resolution in Council or Committee, he found many opportunities of showing his mastery of that style of composition. I remember one occasion on which he made use of this acquisition with much effect. The question under discussion in the Council was the serious cost of the investigations which the Society undertook at the request of the Government. This was naturally a matter wherein the Treasurer was specially concerned, inasmuch as the expenses of these investigations were defrayed out of the Society's income, and sometimes amounted to a considerable sum, while the uncertainty of their probable cost always raised a difficulty in the framing of the budget for each year. He accordingly drafted and read aloud the following statement:—

"Whereas the President and Council have frequently been requested by various Departments of the Government either to advise them upon, or, in some cases, to undertake the supervision of, and in others, the entire responsibility for, various scientific investigations of national importance no provision has been made by Government to meet expenses to which the Society has necessarily been put in acceding to these requests."

It was amusing to hear the emphasis which, as he looked up from his manuscript, he laid on the preposition at the end of each clause. He succeeded in getting negotiations set on foot with the Government, which resulted in a satisfactory arrangement for the future. The Treasurer then prepared a new regulation, approved by the Treasury, whereby adequate provision was made

> "for any expenditure which may be incurred by the Royal Society in undertaking, controlling, supervising, or advising upon matters which the President and Council may, at the request of the Government, undertake, control, supervise, or advise upon."

As he read to the Council this document, which recorded the end of all the troubles of the past, there was a twinkle of quiet satisfaction in his eyes as, with much firmness in his voice, he pronounced each of the four verbs which described the varied kind of work done by the Society for the Government. The very sound of the words seemed to be pleasant to his ears.

Probably at no time in its history had the Royal Society been in close relations with so many Departments of Government as during Kempe's Treasurership. There can be little doubt that his sagacity and clearness of judgment in these conferences were of great value in removing difficulties, and impressing on the official mind the nature and extent of the assistance which the Society could render. It was doubtless in acknowledgment of these services that, in 1912, the honour of Knighthood was conferred upon him. His legal distinction had already been recognised by the University of Durham, which, in 1908, conferred on him the honorary degree of D.C.L.

In the early years of last century when the growing interest in the progress of science was beginning to suggest the creation of independent societies for the prosecution of research in different branches of enquiry, the movement was looked upon with disfavour by some of the leaders of the Royal Society, unless the new organisations were placed under the wing of that Society. This subordination was vigorously resisted, and many such societies have since then been successfully founded, and have been of the greatest value in extending the cultivation of the sciences which they represent. Yet the Royal Society, with the hearty goodwill of these younger associations, retains its time-honoured prestige, and finds that its activities have grown more varied and pressing than ever. Its Treasurership is an office that naturally brings the holder into contact with the other younger scientific bodies, and affords many opportunities for the promotion of friendly intercourse with them. Never were these amenities more happily secured than during Sir Alfred Kempe's tenure of the post. It may be mentioned here that of one of the younger scientific coteries—the Royal Institution—he was a member for half a century, serving five times on its Board of Management, and taking a keen personal interest in its welfare.

A feature that should not be omitted from this sketch of Kempe's life was his abounding love of mountain scenery, which for many years drove him to spend his holidays in Switzerland. The lure for him was not so much the joy of reaching almost inaccessible peaks (though he could wield his ice-axe and take his share of adventurous climbing), as the quiet enjoyment of the grandeur and beauty of the mountain-world, and the pleasure of being once more amidst the Alpine flora. In the gratification of this passion he must have visited the Alps between forty and fifty times. To the last he maintained his keen interest in the literature of mountaineering.

Allusion has already been made to Kempe's love of music. Gifted with a good counter-tenor voice, he early began to sing. Even at school he was a member of the St. Paul's School Choral Society, where he sang the treble parts and later the alto. At Cambridge, among his college friends and fellow students, he gained a musical reputation, and became librarian of the University Musical Society before he was widely known as a mathematician. He had a piano at his rooms, on which he no doubt played the accompaniment to his vocal practisings. One of his friends at the time thus described the relationship between the instrument and its owner:—

Mistress of humble tones and haughty,
Kempe calls me his piano-forte ;
He plays me when a problem fails,
And rises lighter from the scales.

He sang in the Bach Choir under Otto Goldschmidt, who retained a pleasant memory of his "beautiful counter-tenor." He was a member of the Moray Minstrels, a private men's choir of glee-singers, where the peculiar quality of his voice enabled him to sing the alto parts. From the weekly meetings, the rehearsals, and the concerts of this choir he was seldom absent until the association was dissolved in the summer of 1907. He likewise occasionally gave his aid to the Westminster Abbey Choir at the evening service.

From this sketch of his career it must be obvious that for at least the last twenty years of his life Kempe was practically carrying on two professions. The ecclesiastical work of the chancellorships of half-a-dozen dioceses, and ultimately the burden of the great diocese of London, together with his professional engagements as a barrister, would have sufficed, it might be supposed, to keep any man fully employed; but, in addition, he had the serious task of piloting the Royal Society through its financial undertakings, as well as taking an ample share in the conduct of its other general business. He never shrank, however, from the discharge of his many duties. Whether or not it was this accumulation of work that overtaxed his strength, his health broke down in 1917, while the War was still in full strain. He nevertheless maintained a brave fight against increasing weakness, until at last, in the summer of 1919, he felt compelled to resign the Treasurership of the Royal Society; but he consented to retain his seat in the Council.

The intimation of the Treasurer's resignation filled the Royal Society with sorrow as of a personal bereavement and a sense of unlooked-for and almost irreparable loss. These feelings were well expressed by the President, Sir J. J. Thomson, in his Address to the Fellows on the following St. Andrew's Day:—

"It was," he said, "with the greatest regret, almost with consternation, that the Council heard from Sir Alfred Kempe that the state of his health obliged him to resign the office of Treasurer, which he has held for 21 years. It is difficult to find words adequately to express our indebtedness to him. By his sagacity, his long experience of the affairs of the Society, and his legal knowledge, he has rendered invaluable services in our councils and in directing the policy of our Society. He carries with him on his retirement from the office which he has so long and worthily held the thanks and good wishes of every member of the Society."

After his retirement from the Treasurership there appeared for a time the possibility that his life might be prolonged. But at last pneumonia supervened, and he quietly passed away on April 21, 1922. Sir Alfred Kempe was twice married: first in 1877 to a daughter of Sir William Bowman, Bart., M.D., F.R.S., who died in 1893; and secondly in 1897 to the elder daughter of his Honour Judge Meadows White, Q.C., who survives him. By the second marriage there are two sons and one daughter.

It is not easy to describe the personal charm which endeared Sir Alfred Kempe to all who came to know him. His modesty, urbanity and frankness were at once apparent; at the same time his sound sense, and the touch of humour or flash of wit with which he would often enliven a formal conversation, made him singularly attractive. The lasting affection of those who were privileged to enjoy his more intimate friendship was won by his combination of genial qualities, above all by the overflowing kindliness of his nature. His humility of mind and antipathy to anything like self-advertisement read a continual lesson to the ambitious. Thoroughness in all that he undertook was one of his most characteristic virtues. Not less conspicuous was the friendly readiness with which he would put his wide knowledge and experience at the service of others. As scientific circles are not free from the irritability and combativeness that affect other coteries of men, Sir Alfred was again and again appealed to as the irresistible peacemaker. Amid all his various gifts of character there was the glow of his pure Christian soul, which, while never obtruding his religion, could not conceal its benign and dominant influence in his life.

ARCHIBALD GEIKIE.

SIR WILLIAM CHRISTIE, 1845–1922.

William Henry Mahoney Christie was born at Woolwich, on October 1, 1845. His father, Samuel Hunter Christie, was Professor of Mathematics at the Royal Military Academy, and Secretary of the Royal Society from 1837 to 1854, and is remembered for researches in Magnetism and as an independent inventor of Wheatstone's Bridge. Christie was sent to King's College School, London, where the diversity of his studies is indicated by the award to him of a prize for Hebrew. He obtained an open Scholarship in Mathematics at Trinity College, Cambridge, in 1864. He was fourth Wrangler in 1868—the first three places in the Tripos being taken by Lord Moulton, Sir George Darwin, and Mr. C. Smith (afterwards Master of Sidney-Sussex College). He obtained a Fellowship at Trinity the following year, and remained in College for a short time, taking private pupils. As an incident of this period of his life he occasionally recalled with pleasure a reading party he conducted in the long vacation to the West of Ireland, where he enjoyed rowing and shooting with his pupils.

In the autumn of 1870, on the recommendation of Sir George Airy, Christie was appointed Chief Assistant at the Royal Observatory, Greenwich, in succession to Stone, who had been made H.M. Astronomer at the Cape. At that time the astronomical work of the Observatory consisted in regular observations of Sun, Moon, Planets and the brighter stars with the transit-circle, supplemented by observations of the Moon with the Altazimuth; observations of γ Draconis with the Reflex Zenith Tube for determination of aberration and parallax; and occasional observations of planets with the 13-inch Equatorial, and of occultations. The most valuable of these—and its importance is very great—is the regular and continuous determination of positions made with the transit-circle. This instrument, erected by Airy in 1851, had been studied by Airy himself and by Stone, in order to detect and eliminate any source of systematic error. Christie, however, found that a small error had been allowed to creep in, owing to the wear of the micrometer screws of the reading microscopes. New gunmetal screws were obtained, and the possibility of error from this cause in the future was diminished by reversing the direction of alternate microscopes. A more important contribution to the study of the systematic errors of the Greenwich declinations was published by him in 1880. In this paper he discussed (1) the value of the constant of refraction at Greenwich, (2) the correction to be applied to the zenith distances, in view of the systematic discordance between observations made directly and by reflection at the surface of mercury, (3) the value of the Latitude of Greenwich. In his solution of this somewhat involved and indeterminate problem Christie's judgment was correct, and his conclusions were generally concurred in by Newcomb, who reconsidered the question in 1890.

In 1870 neither spectroscopy nor photography was applied at Greenwich to the study of the heavenly bodies, and it was largely due to Christie that they were introduced a few years later. He was elected to the Council of the Royal Astronomical Society in 1872, and so came into touch with Huggins and Warren de la Rue and other astronomers whose interests lay in these newer methods of research. De la Rue pressed on Airy the desirability of continuous photographic observations of Sun Spots, while Huggins advocated spectroscopic observations of Sun and Stars. They had in Christie an enthusiastic ally within the Observatory. In a letter to Airy, in May, 1872, Huggins writes, "I understand Mr. Christie, who is zealous in the matter, to say, that you would be agreeable to this course." Government sanction was obtained for the appointment of an assistant for photographic and spectroscopic observations, and Mr. E. W. Maunder was chosen for the post. Solar photography was begun at Greenwich with an instrument of De la Rue's in 1874. In order to fill up gaps in the daily Greenwich series, the co-operation of the Observatories at Dehra Dûn and Mauritius was secured through the Solar Physics Committee. Four similar instruments constructed by Dallmeyer for photographing the Sun at the transit of Venus were available. It was arranged that the measurement and discussion of the photographs should be made at Greenwich. The continuous record of Sun Spots in position and area, commenced by Christie and Maunder, continues as a valuable part of the routine work of the Observatory, and has formed the basis of valuable discussions on the behaviour of Sun Spots and their relationship to Magnetic Phenomena.

The same success did not attend the spectroscopic researches. A good deal of time and energy was given by Christie and Maunder to the determination of the velocities of stars in the line of sight. The results obtained are now seen to be unreliable. It was not till the introduction of photography by Vogel that results of value were obtained, and not until the inauguration of the Mills Spectrograph of the Lick Observatory in 1894 that thoroughly trustworthy velocities of stars in the line of sight were secured. Nevertheless credit is due to Christie and Maunder for following Huggins in his pioneer work, though none of them solved the difficulties incidental to these refined measurements.

In 1877, while Chief Assistant at Greenwich, Christie founded the "Observatory" magazine. For some years the "Astronomical Register" had dealt with current astronomical topics, but in a somewhat partial manner, and continued to do so till 1886. The "Observatory" has continued for 45 years, as a useful monthly review of Astronomy, in much the same form and character in which it was left by its first Editor.

Christie, who had been an unsuccessful candidate for the post of Director of the Radcliffe Observatory in 1878, and that of the Cape Observatory in 1879, was, in 1881, appointed Astronomer Royal. In Airy's long tenure of office the Observatory had been admirably organized, but at the time of his retirement Astronomy had developed in several directions which were not represented at Greenwich. Christie gradually obtained increase in the equipment of the Observatory and widened its field of work. The first addition to the buildings was an additional computing room, built over one already existing and surmounted by a dome intended to house the photo-heliograph, but in which the Astrographic Equatorial was afterwards installed.

In 1885 he represented to the Admiralty the desirability of increasing the optical means of the Observatory, and received its assent to the purchase of an object glass of 28 inches aperture and 28 feet focal length. The object glass was constructed by Sir Howard Grubb from discs made by Chance and Mantoix. Stokes made the suggestion that an object glass should be constructed which could be used for photography by reversal of the crown lens, with further separation of the lenses to correct for spherical aberration. This proposal was adopted, and a large telescope obtained which could be used alternatively for visual or photographic observation. The Equatorial mounting of the 13-inch Equatorial was adapted for this much larger telescope. Also the drum-shaped dome, which was worn out, was replaced by a new one of Christie's design, so contrived as to get the largest possible telescope in the existing building. The bulging shape of the dome gives a unique and somewhat oriental character to the Observatory Buildings. The telescope and dome were not completed till the end of 1893. A half-prism spectroscope was mounted on it and some attempts made at the determination of velocities in the line of sight. Also a few photographs of the Moon and some double stars were taken with the crown lens reversed. But gradually the instrument came to be employed almost entirely for observations of double stars. Airy had regarded this class of work as unsuitable for a public observatory, but peculiarly fitted for amateurs. Gradually, however, a large number of close double stars have been discovered which are beyond the powers of telescopes possessed by amateurs, and it has become necessary that public observatories should take part in these researches. A year before the completion of the 28-inch Equatorial, double-star observation was begun at Greenwich with the 13-inch Equatorial. It was carried on by Mr. Lewis with the new telescope, and gradually became part of the regular work of the Observatory.

About the same time, a 13-inch photographic refractor was obtained, in order that Greenwich might take part in the international photographic map of the heavens. The instrument was constructed by Sir Howard Grubb and completed in 1890. Work on the Greenwich section of the Chart and Catalogue commenced soon afterwards. Christie took his share with other astronomers in forwarding this project, and attended the meetings at Paris where its details were discussed. He insisted on the importance of the measurement and discussion of the photographs being made at the observatory where they were taken, and not sent to a central bureau. He designed a "duplex" micrometer, in order that the stars on overlapping plates might be automatically identified. The measurement of the plates and the publication of results occupied a number of years and was satisfactorily completed. He realized that the value of this work would be enormously increased by re-observation of reference stars, instead of reliance on positions given in the catalogues of the *Astronomische Gesellschaft*, which referred to an epoch 25 years earlier, and reference stars for the Greenwich and Oxford Zones were observed at Greenwich with the transit-circle. In connection with the Astrographic work he discovered a simple and useful empirical formula connecting the photographic magnitude of a star with its measured diameter. Under Christie's direction the Astrographic Catalogue served as an education of the Greenwich astronomers in astronomical photography. This bore fruit in the more refined work of the determination of the Solar Parallax from observations of the planet Eros, in progress at Greenwich from 1900 to 1908, and in the still more refined work on Stellar Parallax carried on since his retirement.

The most important addition made by Christie to the Royal Observatory was commenced in 1890 and completed 1898. It is a cruciform building designed by Sir Frank Crisp, of the Admiralty, which contains office rooms on the ground floor, libraries and workshop in the basement, and store-rooms for photographs and records on the upper floor. The central portion of the building is used as a store-room, and is surmounted by a 36-ft. dome originally built to house Lassell's 2-ft. mirror presented to the Observatory by the Misses Lassell. Before the building was completed, Sir Henry Thompson generously offered to provide a 26-inch refractor and a 30-inch reflector, both on the same Equatorial mounting. The Equatorial and the refractor were constructed by Sir Howard Grubb. The silver on glass mirror was made by Dr. A. A. Common, who generously gave his own personal services to its manufacture.

The new building has been of the greatest service, as by it suitable provision has been made for the accommodation of the staff of the Observatory and various measuring machines. It is a permanent monument to Christie's foresight in seeing the needs of the Observatory and his persistence in putting them forward. Both the refractor and the reflector have been of great value, and have to a large extent determined the nature of the activities of the Observatory.

About the same time a new Altazimuth was erected according to Christie's designs, to replace Airy's instrument, which did not give results of sufficient accuracy. The new Altazimuth is essentially a transit circle which can be mounted in any azimuth. It usefully supplements the transit-circle for observations of the Moon in first and last quarters.

These additions to the Observatory buildings compelled another change. The amount of iron in the new buildings, especially the domes, made its presence felt on the magnets. A new site for them was secured on a piece of ground in Greenwich Park, which was lent to the Admiralty by the Office of Works. A building free from magnetic material was erected for taking absolute magnetic observations. The meteorological instruments which had gradually been surrounded by buildings were also removed to this more open situation. Christie threw all his energies into the rebuilding of the Observatory and the renewal of the instrumental equipment.

The criticism may be made, in which he himself agreed, that the domes are rather small for the instruments they contain. He was in a very different position from one who is designing an entirely new observatory. The conditions of site and existing instruments and buildings imposed severe limitations. He adapted these as far as possible, and the result of his activity has been, not merely to prevent the Observatory from falling behind, but to provide the means by which it should be kept for a generation or more in the front rank. Concurrently with the increase of instruments was an increase in the staff of the Observatory, and the addition of a second Chief Assistant has proved of great value in maintaining a fresh scientific outlook.

During Christie's tenure of office several expeditions were made by members of the staff for determination of longitude or for observation of Solar Eclipses. He himself went to Japan in 1896, to India in 1898, to Portugal in 1900, and North Africa in 1905, with other astronomers, to observe the Solar Eclipses of those years. He constructed a telephoto combination with a 9-inch object-glass and a 4-inch enlarging-lens, and obtained large-scale photographs of the Corona, except in 1896, when the sky was overcast.

For many years Christie took an important part in the activities of the Royal Astronomical Society. He served on the Council for a continuous period of 41 years, from 1872–1913. He was Secretary from 1880 to 1882, and President, 1888–90. He was elected a Fellow of the Royal Society in 1881, and served on the Council 1883–85, 1889–1901, 1900–01, being a Vice-President in the two latter periods.

The several determinations of the longtitude Paris-Greenwich, the meetings of the Bureau of Standards, and especially of the Astrographic Chart and its offshoot, the determination of the Solar Parallax from observations of the planet Eros, brought opportunities which he welcomed of co-operation and of forming friendship with French and other colleagues. The expedition to Japan gave him the opportunity of visiting Harvard and Yerkes. He was always glad to see foreign astronomers at Greenwich, and took the opportunity of introducing members of his staff to them. His general attitude with regard to the international relations of astronomers, was that great good came from meeting and discussion of problems, but that resolutions binding observatories to common action should only be adopted with great caution, as they might be difficult to carry out in some circumstances and might tend to diminish individual initiative.

Christie received the distinction of C.B. on the occasion of Queen Victoria's Diamond Jubilee, and was promoted K.C.B. in 1904. He was an Honorary D.Sc. of Oxford, and Corresponding Member of the Academy of Sciences of Paris, the Imperial Academy of Sciences of Petrograd, and of the Italian Spectroscopic Society.

In 1881 he married Violette Mary, daughter of Sir Alfred Hickman, of Wolverhampton. Mrs. Christie died in 1888, leaving two sons, one of whom died in childhood. His elder son, Mr. Harold Christie, lived at the Observatory with his father until 1910. On October 1, 1910, Christie retired with the good wishes of his staff, and he and his son went to live first at Woldingham and afterwards at Downe in Kent.

Soon after his retirement he was nominated a member of the Board of Visitors of the Royal Observatory, and attended the Annual Meeting regularly, besides occasional visits. He also kept in touch with his astronomical friends by frequent attendance at the meetings of the Royal Astronomical Society.

In the winter months he sometimes went abroad; in 1921 he visited Jamaica and paid a visit to his friends, Mr. and Mrs. W. H. Pickering, at the Observatory at Mandeville. In 1922 he started for Mogador a few days after meeting his friends at the Royal Astronomical Society Club in January. He seemed in fair health, but was taken ill suddenly and died on board ship on January 22. He was buried at sea shortly before the ship reached Gibraltar.

F. W. D.

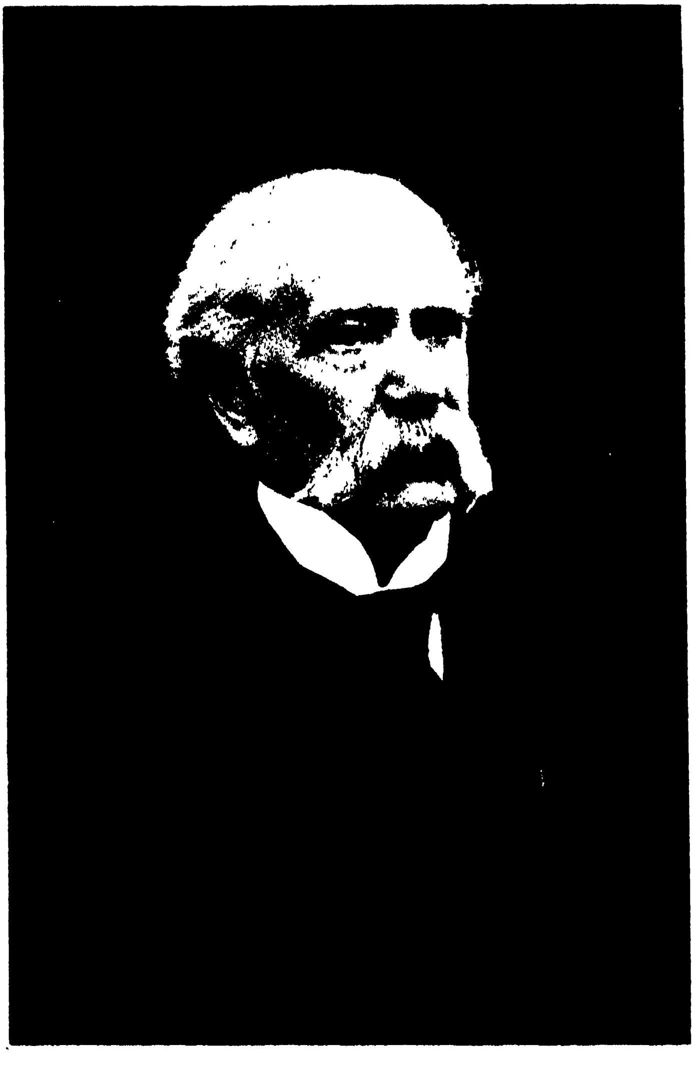
WILLIAM GOWLAND, 1842-1922.

WILLIAM GOWLAND was born in Sunderland on December 16, 1842, the son of George Thompson and Catherine Gowland. Originally intended for the medical profession, he worked with a doctor at Sheffield for two or three years. Purely scientific pursuits, however, attracted his attention, and in 1868 he became a student at the Royal College of Chemistry. In the succeeding two years he secured the Associateship in Mining and Metallurgy at the Royal School of Mines and was awarded the Murchison Medal in Geology and the De la Beche Medal in Mining.

During 1870-2 he obtained employment as chemist and metallurgist to the Broughton Copper Company, Manchester. In 1872 he went to Japan and for several years he worked as chemist and metallurgist to the Imperial Mint at Osaka. In 1878 he became assayer and chief of the foreign staff and adviser to the Imperial Arsenal. During this time he made several expeditions into the mountains besides a journey through Korea, where he carried on work for the Japanese Government. On his return to England, in 1889, he received the order of the Rising Sun, with which he was personally invested by the Emperor of Japan. In 1890 he returned to the Broughton Copper Company as chief metallurgist.

In the same year he married Joanna, youngest daughter of the late Murdoch Macaulay, J.P., of Linchader, Isle of Lewis. She died in 1909.

In 1894 he undertook the duties of Examiner in Metallurgy to the Board of Education at the same time that he acted as External Examiner to the



W. Gowland

Royal School of Mines. In 1902 Sir William Roberts-Austen, Professor of Metallurgy in the Royal School of Mines, died and Gowland was appointed his successor. It was a most fortunate circumstance that Gowland was available for the vacant post, as he brought to the school his very broad practical experience acquired in Japan as well as this country. He at once began to introduce modern methods of examining steel, especially the use of the microscope. A very genial colleague, he became immediately very popular both with staff and students. Unfortunately, the Civil Service age rule compelled his retirement in 1909. He was, however, invited by the Governors to return in 1913.

In 1910 he married Maude Margaret, eldest daughter of the late D. J. Connacher, who, with one daughter of his first wife, survives him.

Gowland wrote frequently on metallurgical subjects, but chiefly in connection with his observations in Japan. His first paper recorded in the Royal Society Catalogue appeared in the 'Chemical News' (1891), on "Native Copper from Yunnan, China." In the 'Journal of the Society of Chemical Industry' (vol. 13, pp. 463–470) there is a paper on "A Japanese pseudo-Speise (Shirome) and its Relation to the Purity of Japanese Copper and the Presence of Arsenic in Japanese Bronze." Another (vol. 15, pp. 404–413, 1896) is entitled "Japanese Metallurgy. Part I.—Gold and Silver and their Alloys." Other papers occur in the 'Smithsonian Report' (pp. 609–661, 1894), "On the Art of Casting Bronze in Japan," and in the 'Journal of the Chemical Society' (vol. 51, pp. 410–416, 1887), "On Silver containing Bismuth."

But Gowland was not only a metallurgist, he was more especially an ardent and enthusiastic antiquarian, and his contributions to 'Archæologia' and to the 'Proceedings of the Society of Antiquaries' and the 'Journal of the Anthropological Institute' were all of great interest and importance.

The following is believed to be a complete list of these publications:—

From Archæologia—

Vol. 55, pp. 439–524, 1897. "The Dolmens and Burial Mounds in Japan."

Vol. 56, pp. 13–20, 1899. "Analysis of Metal Vessels found at Appleshaw, Hants, and of some other Specimens of Roman Pewter."

Vol. 56, pp. 267–322, 1899. "The Early Metallurgy of Copper, Tin, and Iron in Europe, as illustrated hy Ancient Remains, and the Primitive Processes surviving in Japan."

Vol. 57, pp. 113–124, 1900. "Remains of a Roman Silver Refinery at Silchester."

Vol. 57, pp. 359–422, 1901. "The Early Metallurgy of Silver and Lead. Part I.—Lead."

Vol. 58, pp. 37–105, 1902. "Recent Excavations at Stonehenge."

Vol. 69, pp. 121–160, 1918. "Silver in Roman and Earlier Times. Part I.—Pre-historic and Proto-historic Times."

From the 'Proceedings of the Society of Antiquaries'—

(2) Vol. 16, pp. 330–334, 1897. "On the Composition of Bronze, Copper, etc., in the Hoards found at Grays, Thurrock, and Southall, and on Experiments on the Manufacture of Ancient Bronze."

(2) Vol. 20, pp. 194, 1905. "Note on Iron Bars used as Currency."

(2) Vol. 20, pp. 242–245, 1905. "On some Crucibles from Rhodesia."

(2) Vol. 21, pp. 20–21, 1905. "Note on Leaden Grave Crosses found in London."

Report of the Research Committee of the Society of Antiquaries—

No. 3. Excavations at Hengistbury Head, Hampshire, in 1911–12. Appendix II (By Prof. W. Gowland). "Report on the Metals and Metallurgical Remains from the Excavations at Hengistbury Head." Pp. 72–83.

Journal of the Anthropological Institute.

Vol. 24, pp. 316–330, 1895. "Notes on the Dolmens and other Antiquities of Korea."

Vol. 36, pp. 11–38, 1906. "Copper and its Alloys in Prehistoric Times."

Vol. 37, pp. 10–46, 1907. "The Burial Mounds and Dolmens of the early Emperors of Japan."

Vol. 42, pp. 235–287, 1912. "The Metals in Antiquity."

The first May lecture delivered (1910) before the Institute of Metals on "The Art of Working Metals in Japan;" Presidential Address to the Institute of Metals (1912) on "Copper and its Alloys in Early Times;" "Recent Excavations at Stonehenge," communicated to the Society of Antiquaries in 1901, contains an account of explorations undertaken by a committee appointed jointly by the Society of Antiquaries, the Wiltshire Archæological Society, and the Society for the Protection of Ancient Buildings, who appointed Prof. Gowland as representative to conduct the exploratory work. The nature and origin of the rock-fragments found in the excavations were reported on by Prof. J. W. Judd.

Numerous stone axes and hammers, stones probably of neolithic age, were found. In the surface layers were a few coins of Roman and more recent date. Gowland came to conclusions concerning the origin of Stonehenge fundamentally different from those which had been generally accepted. In view of the large accumulation of chips and fragments of the Sarsen and "blue stones" of the monolith immediately upon the site of the circle, and from other considerations, he was led to believe that the stones were not brought from a distance but were originally "grey wethers" lying on the surface of the chalk downs, and that the stones were selected at no great distance from the spot where the structure stands, and were trimmed at the spots where they were found. His own words as to this conclusion are as follows: "I would therefore suggest as probable that when the early inhabitants of this island commenced the erection of Stonehenge, Salisbury Plain was sprinkled over thickly with the great white masses of the Sarsen stones (grey wethers) and much more sparingly with the darker coloured boulders (the so-called, 'bluestones'), the last relics of the glacial drift circle had been nearly denuded away. From these two kinds of materials the stones suitable for the contemplated temple were selected. It is even possible that the abundance and association of these two kinds of materials, so strikingly contrasted in colour and appearance, at a particular spot, may not only have decided the size, but to some extent have suggested the architectural features of the noble structure of Stonehenge."

Gowland was always actively interested in the history and application of metals, and naturally his long residence in Japan afforded many opportunities of enquiring into the early methods employed in that country. These were described in the many interesting and instructive papers and addresses he has left behind, and of which a list has been given.

He was an active member of many Societies, among the rest the Society of Antiquaries, the Chemical Society, of which he became a Fellow in 1871, the Institute of Chemistry, the Iron and Steel Institute, the Society of Chemical Industry, the Royal Anthropological Institute, of which he became President in 1905–7, and of the Institution of Mining and Metallurgy, of which he was President 1907–8, as also of the Institute of Metals in 1912. He was elected into the Royal Society in 1903, and served on the Council in 1912–14.

As a young man Gowland was very fond of rowing, and he was the first to introduce the sport into Japan. In order to encourage boat racing among the Staff at the Mint he had two modern " eights " built, but the Japanese found them too unstable and preferred two of their own boats which they presented for Prof. Gowland's inspection. To his astonishment he found that they had selected a pair of cutters, and had fitted them with port and starboard lights!

Since November last Gowland's health began to fail, but he was still able to work at the fourth edition of his book on the "Metallurgy of the Non-Ferrous Metals." His illness, however, took gradually a more serious turn, and on Whit Sunday he became suddenly paralysed and passed away on June 10.

W. A. T.

E. H. GROVE-HILLS, 1864–1922.

Edmond Herbert Hills was born on August 1, 1864, at High Head Castle, Cumberland. His father had been Judge of the Court of Appeal in Egypt, and his mother was a daughter of Sir William Grove, well known both as a judge and as a man of science. Educated at Winchester and the Royal Military Academy, he obtained a commission in the Royal Engineers in 1884. From 1884 to 1886 he was at the School of Military Engineering, at Chatham, and in regular course went to Gibraltar in 1887. He obtained his captaincy in 1893.

In 1893 he went with Sir Edward Thorpe to observe the eclipse of the sun on April 16 at Fundium, in Senegambia. Hills took two slit spectroscopes, with which he obtained successful photographs of the spectrum of the prominences and corona. He published a report on his observations in the 'Proceedings of the Royal Society,' 1894. Attention may be drawn to the number of the hydrogen lines photographed. In 1894 he contributed to the Royal Astronomical Society an account of a simple method for obtaining longitudes in the field from photographs of the moon. His interest in eclipses took him to Japan in 1896, but owing to cloudy weather no results were obtained. In 1898 he had better fortune in India and obtained excellent photographs of what was then called the "flash" spectrum, but is now generally known as the chromospheric spectrum. These photographs and those taken at the same eclipse by Sir Norman Lockyer's expedition demonstrated—what was till then a vexed question—that the spectrum of the chromosphere was not a reversal of the ordinary solar spectrum. Hills retained his interest in eclipse work, and was for many years secretary of the Joint Permanent Eclipse Committee. His spectroscopic equipment was lent to observers of subsequent eclipses and was afterwards presented by him to the Cambridge University Observatory.

From 1896 to 1899 Hills was Instructor in Chemistry and Photography at the School of Military Engineering at Chatham, and from 1899 to 1905 was in charge of the Geographical Section of the General Staff at the War Office. During part of this time he was responsible for the production and supply of maps required in the Boer War. This was not an easy task, as he found the country had been very imperfectly surveyed. In 1904 the Intelligence Section of the War Office was re-organised, and Hills pressed successfully for the enlargement of the staff of the Geographical Section. The advantages of this reorganisation were reaped ten years later, at the outbreak of the war. During this period he was brought into contact with several geodetic questions, particularly Sir David Gill's project for measuring a long arc of meridian from South Africa to Cairo. To this scheme, which was carried out to a considerable extent, Hills gave hearty support.

He was Secretary to the Commission of which Sir T. H. Holdich was

President, appointed in 1902 to delimit the boundary between Chile and Argentina. In 1901 these countries were on the verge of war, but in December agreed to submit their differences to the arbitration of King Edward. The difficult question of fixing a boundary on the watershed of the snowy Cordilleras was settled to the satisfaction of both countries. For his services in this Commission Hills was awarded C.M.G.

On the termination of his period of service on the General Staff in 1905, Hills retired from the active list, and contested Portsmouth on the Unionist side. He was unsuccessful, and did not pursue active participation in politics any further. While at the War Office he gave close attention to Imperial surveys, and, after he left, his ability and experience were utilised in many parts of the British Empire. He inspected and reported on the survey work for the Canadian Government. He was appointed by the Secretary of State for the Colonies to report on the Survey Departments of British East Africa and Uganda. After completion of this work he made a similar inspection in Ceylon. In these visits he entered sympathetically into the various difficulties experienced in the different countries, and by his helpful advice materially improved the standard of survey work in the British Empire.

A subject in which Hills was keenly interested was the Variation of Latitude. In co-operation with Sir Joseph Larmor he derived from the curves of the movement of the pole, given by observation at the international stations, the torques which gave rise to these displacements, and endeavoured to find the dynamical causes of the movement of the pole. His interest in this subject led him, in another direction, to design and have constructed a photographic instrument in which the position of the vertical is given by suspension of the instrument. The trials of this instrument were in progress at the time of his death.

At the outbreak of war he had started for Russia to observe the eclipse of August, 1914, but was recalled and appointed Assistant Chief Engineer on the Eastern Command. He reached the temporary rank of Brigadier-General and, in 1918, was awarded the C.B.E.

After the war he resumed his scientific activities. He filled the important post of Honorary Secretary of the Royal Institution, for which he was admirably fitted by his wide knowledge and interests. He accepted the post of Honorary Treasurer of the Royal Astronomical Society in 1922, and zealously set to work to improve the financial position of the Society, whose activities were in some danger of being impaired by the great increase of prices, especially of printing. By his efforts a considerable sum of money was obtained, and a large accession of new Fellows to the Society was secured. His life-long interest in this Society is perpetuated by a legacy to it of a valuable library of old works on Astronomy which he had collected.

Hills married, in 1892, Juliet, daughter of James Spencer-Bell, M.P. They had three children, two sons and a daughter. Both sons volunteered for the air-service and gave their lives in the war. This heavy bereavement was borne bravely, and no difference was seen in public in the discharge of his military duties during the war, or in his services to the scientific bodies with which he was connected.

Colonel Hills, who recently prefixed an additional surname, Grove, on becoming heir to his uncle, General Sir Coleridge Grove, inherited many of the gifts of his ancestors. He was endowed with a wonderful memory, and a mind which quickly grasped the essentials of any problem which arose. These qualities served to make him an excellent administrator as well as a scientific investigator. His recreations were shooting and fishing, and in both of these arts he excelled.

In May, 1922, he attended the meetings of the International Unions of Astronomy and Geophysics in Rome, and in June took part in the celebration of the centenary of the Royal Astronomical Society. He was suffering from what was thought at the time to be only temporary indisposition. Soon afterwards it became clear that the illness was of a very serious nature. It made rapid progress and he died on October 2, 1922, to the great grief of a large circle of friends. F. W. D.

G. W. WALKER, 1874–1921.

By the death of Mr. George Walker Walker, the Royal Society has lost a physicist of originality gifted alike with mathematical and experimental aptitude.

Mr. Walker was the only son of Mr. John Walker, of Aberdeen, and was born on February 24, 1874. He was a foundationer at the Robert Gordon's College, Aberdeen, and he started his working life as a practical engineer. He then obtained an appointment with Messrs. C. and P. H. Chalmers, of Aberdeen, where he remained for one year. His interests in science, however, were so great that he continued its study in the evening classes of Gordon's College, and with such success that he obtained a national scholarship which brought him to South Kensington. Here in due time he obtained his associateship. While at South Kensington he greatly impressed Sir Arthur Rücker with his mathematical ability. In addition to his theoretical and experimental work under Rücker, he followed closely all that was being done at the time with quartz fibres, with which he became practically familiar. Rücker urged him to go to Cambridge, where he obtained a Sizarship at Trinity College. He was Fourth Wrangler in 1897 and Smith's Prizeman and Isaac Newton Student in 1899, and became a Fellow of Trinity in 1900. In 1901 he studied at Göttingen. From

1903 to 1908 he was Lecturer in Physics at Glasgow University. He was then appointed Superintendent of Eskdalemuir Observatory, for which he was so well qualified. Here he remained four years. From 1912 to 1915 he was engaged on the New Magnetic Survey of the British Isles. He was elected a Fellow of the Society in 1913. He assisted the Earl of Berkeley in his laboratory at Boar's Hill in 1915, and he became Halley Lecturer in 1916.

During the period 1898—1914 Mr. Walker published, in the 'Quarterly Journal of Mathematics,' the 'Proceedings of the Royal Society,' and the 'Proceedings of the Physical Society,' a number of original mathematical papers, dealing chiefly with the atomic constitution and the associated electrical and optical phenomena of gases, with some experimental confirmation.

In May, 1918, Mr. Walker was appointed chief scientific worker at the Royal Naval Mining School, Portsmouth, where the numerous and urgent problems connected with marine mines and detection of submarines afforded abundant scope for his genius. His success in dealing with these problems was very highly appreciated by the Superintendent of Mining and other naval officers concerned, and also by a small Committee of Fellows of this Society, by whom certain of these problems were discussed. Mr. Walker's work called for high mathematical and experimental skill, as well as capacity for invention, and in none of these was he found wanting. This work is obviously confidential. In November, 1920, while engaged on experimental work at Falmouth, he contracted a chill, which developed into lung trouble, and was the cause of an abscess, for which operation became necessary. He went into University College Hospital, Gower Street, in July, 1921, where the shock due to two operations led to his death.

Besides his work on the Magnetic Survey, Mr. Walker was an authority on Seismology, being also a warm admirer of Prince Galitzin.

In 1904 he married the daughter of Mr. Gifford, of Aberdeen, who, with one son, survives him.

Dr. Chree and Prof. Turner, who are intimately acquainted with Mr. Walker's investigations on Magnetics and Seismology, have written the appended notices.

C. V. B.

Eskdalemuir Observatory being primarily devoted to the study of terrestrial magnetism, Walker, on his appointment in 1908 as first superintendent, had his attention called to the subject. The regular procedure of a magnetic observatory, the registration of the magnetic elements by instruments which must not be interrupted for experimental purposes, and the use at regular intervals, in a prescribed way, of absolute instruments, the invariability of which is the first consideration, was not altogether congenial to a man of Walker's temperament. In 1910, however, after the transfer of the observatory from the National Physical Laboratory to the Meteorological Office, it was decided, on the advice of the Gassiot Committee, to alter the magnetographs which recorded declination and horizontal force, so that in future they should record the north and west components of force. This alteration did not tend to simplicity, and to most superintendents would not have been welcome, but it afforded Walker an opportunity, which he appreciated, for making an excursion into less beaten tracks.

At a later epoch, when engaged in the magnetic survey of the British Isles he invented a portable variometer, having a quartz fibre suspension, with which he made a number of observations in various parts of Ireland. Only an elaborate set of comparisons at a fixed observatory could decide whether Walker was too optimistic in believing this instrument to rival in accuracy the Kew pattern magnetometer, but it would certainly seem to be a convenient instrument for a detailed study of local disturbances.

Two papers by Walker on the regular diurnal variation of the magnetic elements appeared in the 'Proceedings.' He entertained doubts whether the usual practice of analysing the diurnal inequality in a Fourier series was of much use, and sought to replace the series by a finite formula. He also made an independent enquiry as to how far the potential suggested by Sir Arthur Schuster gave a satisfactory representation of the diurnal variation at a series of stations. He did not consider the agreement between the actual observations and the theoretical expressions very satisfactory, and proposed the addition to the potential of several terms for which he found some physical basis. He concluded, however, that the accuracy of the observational data available was too uncertain to admit the attainment of a decisive result.

Walker's most important contribution to terrestrial magnetism was his magnetic resurvey of the British Isles for the epoch January 1, 1915. This received the combined support of the Royal Society, the British Association, and the Survey Department. The results appeared in the 'Philosophical Transactions,' A, vol. 219. When a complete satisfactory survey has once been made, all that is absolutely necessary is a knowledge of the secular change of the magnetic elements throughout the area. The simplest and least expensive way of obtaining this information, when circumstances allow, is to utilise the mean annual values published by observatories. This method has been employed by A. Angot in France. If the observatories are well distributed, a comparatively small number suffices. The British Isles, however, are surrounded by sea areas, from which no reliable information is forthcoming; and in 1891, the date of Rücker and Thorpe's survey, there was no magnetic observatory in either Scotland or Ireland. Thus so far at least as these countries were concerned, the only possible plan was to re-observe at some of Rücker and Thorpe's field stations. In view of the uncertainties of field observations, the number of stations to be re-occupied to obtain satisfactory measures of secular change is necessarily considerable. Of the 200 stations originally contemplated, Walker occupied 183 during the summers of 1914 and 1915. The remaining stations, situated in the Hebrides, the Isle of Man, and the Channel Islands had to be omitted owing to the war. In his discussion of the results, Walker devoted considerable attention to local disturbances, especially those in the neighbourhood of Melton Mowbray, first detected by Rücker and Thorpe. He brought the subject to the notice of the Council of the Royal Society, who referred it to the Council of the Conjoint Board of Scientific Societies, and a committee was appointed to consider the bearing of magnetic disturbances on the possible occurrence of iron ore. The results of the enquiry thus originated appeared in the 'Philosophical Transactions' as an appendix to Walker's paper. C. C.

G. W. Walker is perhaps best known to Seismologists by the admirable monograph on their subject which he published in 1913 ("Modern Seismology," by G. W. Walker, Longmans), and which has since been the standard work on the subject for the English-speaking world. The book, however, would scarcely have been written (as he tells us in the introduction) but for his unique experience in dealing with the chief varieties of seismographs, all assembled under his care at the Eskdalemuir Observatory, viz., those of Milne, Wiechert, Omori and Galitzin. Walker had a skilful and sympathetic touch with instruments which brought out their good points almost as though they were living things capable of affection; while an injury to them, or even rough handling by incompetent people, hurt him as though they really suffered pain. His interest in them was also catholic. While he appreciated fully the fundamental improvements introduced by Galitzin, he did not lose admiration for the simple pioneer Milne instrument. His book opens with a touching tribute to Milne, "truly the father of modern Seismology," who died just at the time the book was finished, and a year or two later (1916, May 17) he had to mourn the loss of Galitzin, also a close personal friend.

The comparative study of these instruments suggested to Walker designs of his own on different lines, but he was never able to put them into actual shape. He was much impressed by the line of thought followed by Galitzin in a paper published just before his death in the 'Bull. Acad. Sci., Petrograd,' 1915, which (together with another by the same eminent writer) was translated from the Russian on Walker's initiative and published in the 'Roy. Soc. Proc.,' A, vol. 95 (1918–19), p. 492. Galitzin's suggestions were for the direct measurement of accelerations, rather than velocities or displacements, and quite possibly they turned Walker's thoughts in a new direction. During the War he arranged an apparatus of a quite different kind for detecting the direction of an explosion by watching a small bubble of mercury. But his early death has cut short any development of his own or other ideas. To prevent misconception, it should perhaps be remarked that his skill with instruments was not merely that of a good manipulator; his profound mathematical knowledge of the guiding principles is well known, and is clearly manifested in his "Modern Seismology."

He had a special admiration for Galitzin's method of inferring, from the records at a single station, not only the distance of an earthquake centre, but its direction in azimuth; and the suggestion that the epicentre might be located by the azimuth observations only (made at two or more different stations) was probably due to Walker, though it was published in the joint names of Galitzin and Walker ('Nature,' August, 1912). In his "Modern Seismology," p. 66, Walker writes :—

> "The advantages of this method are that it is quite independent of (1) the time at the two stations, and (2) the determination of S, and thus free from any error that attaches to the empirical time curves. It should thus prove of great value in improving the empirical time curves, more especially for short distances, where the influence of finite depth of focus is considerable. For this reason I consider that an instrument which would give the azimuth directly would be of great service even if the remaining portion of the seismogram had to be sacrificed."

It seems probable that we see here the origin of his work with the mercury bubble for detecting the azimuth of an explosion.

His admiration for all Galitzin's work led him to draw attention more than once to the Pulkovo determinations of the angle at which earthquake waves emerge at different stations. These determinations show a minimum value at about 36° from the epicentre, which suggests (if there be no unsuspected error) that the depth of the original phenomenon is considerable, say one-fifth of the Earth's radius. Walker attempted to reconcile other observed facts with this hypothesis of great depth, especially in a paper read soon before his death ('Phil. Trans.,' A, 222, pp. 45–56). He did not meet with any great success in this courageous attempt, but his persistent endeavour indicates his faith in the Pulkovo observations, and his conviction that some explanation must be found for them even by the rejection of ideas already accepted.

H. H. T.

C. G. KNOTT, 1856-1922.

CARGILL GILSTON KNOTT was born on June 30, 1856, at Penicuik, in Midlothian. At the age of 16 he went to the University of Edinburgh, an age customary at that period, and one that had much to recommend it. He became a student under P. G. Tait and worked in his laboratory. This laboratory has been described as an ill-equipped attic, where, without preliminaries, the students were engaged at once upon research, any research that at the moment occupied their professor's thoughts. With the right professor and the right student it proved an extraordinarily stimulating method. Knott, like many other of Tait's students, never lost his love for research. His contributions to science amounted in the end to more than eighty papers, great and small. His loyalty to Tait would have been called veneration if it had not been so evidently an act of affection. The "Life" of Tait, which he subsequently published—Camb. Univ. Press, 1911—is one of the most interesting scientific biographies in the language.

He acted as Tait's assistant from 1879 to 1883, when he was appointed Professor of Physics in the Imperial University of Japan. In this appointment he succeeded an Edinburgh fellow-student—now Sir J. A. Ewing. The period passed in Japan by Ewing and Knott, with Milne, had the very important consequence of calling the science of seismology into life, a science whose inception Knott was fond of tracing to an Edinburgh professor, J. D. Forbes, who first constructed and named a seismometer in 1841. Ewing's interests were afterwards diverted, Milne conceived and organised his world-wide survey, and devoted his whole energies to shaping its problems. Knott continued throughout his life to contribute papers on the dynamical aspects. His last important memoir, 'Proc. R.S.E.,' vol. 39, (1919), is a valuable discussion of the accepted times of transmission of earthquake waves over large arcs of the earth's surface, with their interpretation as to the paths followed by the waves, and the light they throw on the elastic state interior to the earth. In a more general sphere Knott was a ready and agreeable writer. He contributed several encyclopædic articles. He published, in 1908, a handbook, 'The Physics of Earthquake Phenomena,' which made the then rapidly-growing subject widely known, and is still to be ranked as a valuable work, one of the best of its kind.

Next to seismology, the bulk of Knott's studies were devoted to magnetism, especially ferro-magnetism. While in the East, in 1887, he carried out, in association with Tanakadate, a magnetic survey of Japan. For this, among other services, he was decorated, before his return, with the Order of the Rising Sun (Fourth Class), an order which he used punctually to wear on occasions of ceremony. He returned to Edinburgh in 1891, where he held in the University successively the post of Lecturer and Reader in Applied Mathematics. He was also Official Adviser to students of science, and to those taking the honours degree in Mathematics and Physics. His kindly, painstaking, cheerful character peculiarly fitted him for such work.

From Tait, Knott imbibed a zeal for quaternions, amounting nearly to a passion. He would embark on fierce controversies about its merits. It cannot be said, however, that he, any more than his master, succeeded in establishing it in favour as a general method of physical research. In strict sequence of tradition, Knott's interest in the Royal Society of Edinburgh was always very active. He became a Fellow in 1880. Most of his writings are found in its publications. He received the award of its Keith Prize (period 1893–5) for his work on magnetic strains in iron and nickel. In 1912, he became General Secretary of the Society, succeeding Chrystal, who had succeeded Tait. In this post his wide genial sympathies found in a new sphere the same scope that they had found in the duties of Official Adviser at the University. The Napier Tercentenary celebration fell in 1914. The chief share of the organisation fell upon Knott, and he edited the Memorial Volume of writings contributed by those who took part in it. In the same capacity he edited for the Society the Collected Papers of Dr. John Aitkens, F.R.S., which were in the press at the time of his death.

Dr. Knott held the degrees of D.Sc. of Edinburgh and honorary LL.D. of St. Andrews. He was elected F.R.S. in 1920. His death occurred on October 26, 1922, after a few hours' illness. He was at work, in his usual health, during the day, and died in the course of the night. He had a happy life, which he owed more than anything else to imperturbable good nature, and I think he can never have had an enemy.

R. A. S.

JACOBUS CORNELIUS KAPTEYN, 1851–1922.

WHEN an astronomer turns away from the solar system to the more remote objects of the heavens a wide field of research lies before him. He may pass from one star to another, finding for each the facts of its distance, true brightness, chemical constitution, temperature, motion in space; some stars will detain attention by additional peculiarities—double stars with orbits to be investigated and masses to be calculated, or variable stars with light-curves of all varieties, or novæ with still more startling changes to be followed. But when all this knowledge of individual stars has been accumulated, there lies beyond a still vaster problem of the structure and organisation of the *system of the stars.* How far does the system extend? How are the different types of stars proportioned in this system? How are their movements distributed and controlled? Is it indeed to be thought of as one organisation at a definite stage of evolution? Or, is it a heterogeneous assemblage of new and ancient formations, the individual stars dying and renewed in endless cycle as fortuitous collision determines? It was this problem which, about the year 1900, took on a new phase, presenting ramifications undreamt of; those who were occupied with sidereal astronomy became conscious of new aims and far-reaching possibilities. The pioneer and most conspicuous leader in the new movement was Jacobus Cornelius Kapteyn.

The year of Kapteyn's death is the centenary of the death of Sir William Herschel, and we have to go back to Herschel for the beginnings of that study to which Kapteyn gave a new impetus. Of Herschel's researches on the system of the stars two main results survive: firstly, his detection of the motion of the sun with respect to this system and approximate calculation of its direction; secondly, his study of the extent of the system by counts of stars, showing that it was much more extensive in the plane of the Milky Way than at right angles to it. For nearly a century efforts were practically limited to improving and elaborating these two results. The problem of the solar motion continued to attract an attention which now seems out of all proportion to its intrinsic interest; but the baffling discrepancies in the various determinations made when the data became more abundant gave a piquancy to the problem. The problem of star-counts was advanced considerably by Seeliger's mathematical investigations. But outside these two classical problems scarcely any fruitful line of investigation was developed. The search for a central sun failed completely. The detection of common proper motion in certain parts of the sky was, however, a noteworthy exception, and it gave a hint of greater things to come.

The state of sidereal astronomy at the opening of the twentieth century is recorded for us authoritatively in Newcomb's book, 'The Stars: A Study of the Universe' (1902). In the last two chapters we find that investigation of the stellar system is beginning to take a wider scope; and here the references are continually to Kapteyn. "The principal steps in this study (statistical study) have been taken by Kapteyn who, in several papers published during the past ten years, has shown how important conclusions may be drawn in this way." Thus we read that Kapteyn has found a remarkable relation between the spectral types of stars and their proper motions, those of large proper motion being almost invariably of Type II. Kapteyn has found that the average motion of a star in space is about 1·8 times the speed of the solar motion. Kapteyn has found (by studying proper motions) the average parallax for the stars of each magnitude, and has computed the whole number of stars per unit volume in different parts of the system. These were some preliminary results of a carefully planned investigation, which aimed at obtaining exact knowledge of the general features of stellar distribution—distribution according to luminosity, according to speed, and according to distance from the centre of the system. It involved great labour in sifting, correcting, and adapting the observational material, and in some cases special efforts to fill the gaps where no observations yet existed. Whilst still in the early stages, this labour was rewarded by an unforeseen discovery, rather disturbing to the main plan, but of the utmost importance to the progress of astronomy. In 1904, Kapteyn discovered that the stars were not moving indiscriminately in all directions, but tended especially towards two favoured directions of motion. The "two star-streams" revealed a complexity of organisation which provoked inquiry; one inquiry led to another, and more and more interesting ramifications of the problem were revealed. The finding of the two star-streams may be said to mark the beginning of the modern era of statistical investigations of the stars. There were no doubt other causes which contributed to the rapid development of the subject. Within the next few years notable additions were made to the data which were so much needed; photographic measures of stellar parallax began to reach the modern standard of accuracy; the first lists of spectroscopic radial velocities were published; Lewis Boss's catalogue of 6,000 proper motions appeared. This wealth of data provided the first essential for progress. But it was Kapteyn's discovery which revealed the great possibilities to be expected from statistical studies, and directed the general attention of astronomers to a practically new field of research. In many cases, too, it must have directly stimulated the collection of the required data.

As Kapteyn was a pioneer, so he remained a leader in the new movement in sidereal astronomy. He had an equal share with Campbell in establishing the next notable result—the dependence of the linear velocity of a star on its spectral type; and he was also prominent in deriving a still more general connection between velocity and luminosity, which has opened up a subject of great possibilities, at present only dimly perceived. He continued to study the distribution of the stars in luminosity and in distance, and obtained results which have been generally accepted as giving a correct idea of the general nature of the system. In late years he took as his own special field the investigation of the helium stars, which are the most massive and hottest of the stars; we owe to him the greater part of our knowledge of the distribution and luminosities of this spectral type.

This is a general outline of Kapteyn's achievement in astronomy; we turn now to the arduous apprenticeship which preceded it. He was born on January 19, 1851, and was educated at the University of Utrecht. After two years spent as Assistant at Leiden Observatory, he became Professor of Astronomy and Theoretical Mechanics at Groningen; he continued to occupy this post until he reached the age of retirement in 1921. Unable to obtain instrumental equipment at his own University, he sought co-operation with those more fortunately provided. At the end of 1885 there began his partnership with Sir David Gill in the Cape Photographic Durchmusterung, by which Kapteyn took over the whole labour of measurement of the plates and the reduction and control of the results. For the northern hemisphere Argelander's Bonn Durchmusterung is the standard register, recording the stars with approximate positions and magnitudes for purposes of identification and nomenclature. This was extended by Schönfeld as far as 23° S. The rest of the sky to the South Pole is now provided for by the Cape Durchmusterung. It reaches a limiting magnitude about 0·5 fainter than the northern Durchmusterungs; moreover, being based on photographic instead of visual observations, it has the advantage that stars could not well be accidentally missed. The Cape Durchmusterung contains 454,875 stars between 18° S. and the South Pole, a figure which gives some idea of the stupendous task. These Durchmusterungs will no doubt ultimately be superseded by the 'International Catalogue and Chart' which extends several magnitudes further and gives the positions with greater precision; but, as Gill foresaw, the date of completion of this enterprise is still distant. An incidental discovery in the course of the work may be mentioned here. Kapteyn and Innes detected a faint star having the enormous proper motion of 8″·7 per annum. This displaced Groombridge 1830 as the fastest star known, and held the record until the discovery of Barnard's star with proper motion 10″·3 in 1916.

For twelve years Kapteyn devoted himself to this heavy task. He had one assistant who remained with him throughout, and a certain amount of irregular and chiefly unskilled help. The measurement was performed with an instrument of his own design, contrived to minimise the labour of conversion of the measures into right ascension and declination. A full account of the precautions taken to render the work trustworthy in every detail is given in Gill's 'History and Description of the Cape Observatory.' He writes—"Probably no work of this kind of like extent has ever been issued so free from typographical and other errors. It is impossible to overestimate its value to southern sidereal astronomy . . . But probably the most valuable result of the C.P.D. to science is the fact that its preparation first directed Kapteyn's mind to the study of the problems of cosmical astronomy, and thus led him to the brilliant researches and discoveries with which his name is now and ever will be associated."

We think that Gill can scarcely have meant to convey that Kapteyn was not already devoted to the problems of the stellar system when he began this great task. Was it not rather an appreciation of the need of the proper motions and other statistics which would result, which impelled him to these self-sacrificing labours? Prior to this he had devoted his energies to measuring parallaxes of stars with the meridian circle; the results can scarcely be considered successful, although he seems to have maintained to the end of his life the belief that with modern improvements the meridian-circle could still do valuable work in this field. In 1889 he was urging wholesale exposures for parallax in connection with the proposed Astrographic Catalogue. But Kapteyn did indeed profit greatly by his work on the Cape Durchmusterung; and if it did not start him on cosmical problems, he became by it more fully immersed in the subject. There is a region lying between purely observational and purely theoretical astronomy in which Kapteyn was unrivalled. We are accustomed to speak of "observational data"; but observations, and especially astronomical observations, are not in general *data*—things given as the solid foundation for theory. They are approximations, probabilities in varying degrees, imperfect, with accidental errors, systematic errors, vicious errors, yet containing some element of truth which can be found and depended on by one who is perfectly acquainted with the ways of them. Only one who has himself had long experience in the technicalities of observational work can appreciate the relative merits of the material offered, how it may be tested and refined. Only one who has a clear theoretical insight can venture to use such imperfect material, realising where the results of his calculations will be dependable, where uncertain, and where they will be sheer extrapolation. To apply the more elegant mathematical solutions we must wait until the data are improved beyond reproach; by more opportunist methods adapted to the heterogeneous material, Kapteyn was able to reach the most important results many years earlier.

His instinct for avoiding the pitfalls contained in imperfect material is well illustrated by his Tables of mean parallaxes of stars for each magnitude. These have an obvious value for theoretical investigations of the stellar system; but in addition they have a regular daily use, and the recent great development of parallax measurement depends on them in a very essential way. It is a very long time since anyone has made a serious attempt to measure the absolute trigonometrical parallax of a star; the necessary refinement can only be secured by differential measures. Consequently it is necessary to apply to each measured parallax a correction for the probable mean parallax of the comparison stars. If we had not been able to estimate this final correction with reasonable accuracy, it would have been almost useless to bring parallax measurement to the degree of refinement reached in modern work. It was Kapteyn's statistical investigations of proper motions which provided the necessary knowledge of the average distances of the comparison stars, and assured us that the differential method of parallax measurement would not break down from this cause. For the average distance of classes of stars the method of parallactic motions gives results with a certainty and accuracy quite unattainable by direct parallax measurement; and mainly by this method the Tables of mean parallaxes were constructed. For the small parallaxes, which form the bulk of modern determinations, the correction to reduce to absolute parallax amounts to a substantial part of the whole.

It is only recently that the accuracy of Kapteyn's mean parallaxes has been fully appreciated. Many astronomers thought at one time that he had made the parallax fall off too rapidly with decreasing brightness. They were misled by ignorance of a certain correlation between the true velocities of stars and their luminosities. Intrinsically faint stars have larger real motions than bright stars; consequently there is a tendency to imagine them nearer than they really are. The discovery of this correlation (which was partly due to Kapteyn) cleared up the discrepancy, and the most recent determinations with improved data show that his original Tables were remarkably near the truth. Another of his Tables, which was also criticised at one time, has been vindicated by the best modern data. This Table gave the number of stars down to each magnitude for varying galactic latitude; and it indicated a very strong galactic concentration of the faintest stars. Some years later investigations were made which disagreed with this; but it now appears that Kapteyn was right.

An example will illustrate Kapteyn's conclusions on the general character of the distribution of stars. Taking a sphere round the sun of radius 560 light-years, he discussed how many stars would be contained. There would be, he said, about 1 star giving 10,000 to 100,000 times the sun's light; 26 from 1000 to 10,000; 1300 from 100 to 1000; 22,000 from 10 to 100; and so on, down to 650,000 stars giving from 1/100 to 1/10 the light of the sun. These are early figures which he himself continually improved on later occasions; but the great thing was that they gave a concrete and substantially correct idea of the nature of the system with which we have to deal. In particular they brought home to astronomers how far from the truth was the idea that apparent brightness was any real clue to distance; and they showed that the selection of stars by brightness (as occurs in our catalogues) is not at all typical of their actual frequency of occurrence in space. This background of ideas is continually in the minds of those who have to discuss the bearing of new results in sidereal astronomy.

The two star-streams were found by Kapteyn in the course of a statistical investigation of the Auwers-Bradley proper motions. His original paper appears in the 'British Association Report' for 1905 (South Africa). When the effect of the solar motion is removed, the two streams must be in opposite directions along the same line, directed towards and away from a point in the sky known as the true vertex. Kapteyn gave the position R.A. 91°, Dec. + 13° for this vertex—which is remarkably close to the definitive position obtained from the best modern proper motions (94·2°, + 11·9°) and from the radial velocities (94·6°, + 12·5°). The direction of star-streaming is precisely in the galactic plane. The details and method of this computation were not published until 1912, when attention was called to it because it was in some respects intermediate between the analytical theories of Schwarzschild and Eddington. But the anomaly in the proper motions, when once pointed out, was very conspicuous, and was confirmed by all those who followed Kapteyn in examining the question. For some years Lewis Boss, a great authority on proper motions, maintained stalwart opposition, but he finally found the evidence of his own excellent catalogue irresistible. The recognition of the two star-streams showed the origin of those discrepancies in the various determinations of the solar motion, which had provoked so much discussion. The most obvious interpretation of the phenomenon is that we have to do with two more or less independent systems of stars which have encountered one another and are at the moment completely mixed; and this was the view which Kapteyn generally took. His own analytical method of treatment, however, was not the embodiment of this special hypothesis, and he was equally ready to accept the interpretation put forward by Schwarzschild in his ellipsoidal theory, by which the star-streaming in preferential directions was simply a feature of the organisation of a single system. There can be little doubt now that the representation by two independent systems conforms more closely to the observed distribution of motions, and that the two "humps" in the velocity-distribution are actually present. But certain differences in the constants of star-streaming for stars of different types and other general considerations lead us still to prefer to seek an explanation of the phenomenon as a stage in the evolution of a single universe on the lines of Schwarzschild's theory.

Kapteyn was unceasing in his efforts to organise the work of obtaining observational statistics, and his advice was greatly appreciated. He always emphasised the great importance of treating the stars in bulk. His view may be illustrated by contrasting the different rates of progress in the determination of spectral type and of radial velocity respectively. Spectral types were being determined at Harvard in thousands, a work necessarily not of the highest refinement, and with many individual errors, but of incalculable value for the statistical discussions which were opening up a general knowledge of the stellar universe. But those who were working on radial velocities seemed to have set their hearts on reaching an accuracy of 1 km. per second, no matter how slow the process; nor could any one of the results be published without long delay, lest haply it should turn out to be an orbital velocity. If we had waited for a corresponding accuracy in proper motions and in parallaxes, our knowledge of the stellar universe would have been meagre indeed. There was certainly at that time (perhaps is still) a field for radial velocity work aiming at an accuracy of only 5 or even 10 km. per second for the bulk of stars; this would have sufficed for many of the more urgent statistical investigations, and would have been of great benefit to progress. But we venture to think that in some of his other proposals Kapteyn was unduly optimistic, and that he was inclined to overestimate the power of statistical methods in sifting truth from error. The scheme for organisation of sidereal research which he had most at heart was his "Plan of Selected Areas." For the fainter stars, numbering many millions, it is unnecessary that every star should be surveyed; and it is important that astronomers should unite in choosing the same samples to examine, so that our knowledge of each sample may be as complete as possible in all branches of investigation. Accordingly, he chose special regions well distributed over the sky, and secured the co-operation of astronomers in making a full survey of these. A Committee was formed in the summer of 1907 to further this programme. It may be hoped that the death of its originator will not cause this valuable work to languish.

Prof. Kapteyn was a frequent visitor to this country, and was in close association with British astronomers, among whom he had many friends. The Royal Astronomical Society awarded him its Gold Medal in 1902. In 1905 he went to South Africa with the British Association. He was elected a Foreign Member of the Royal Society in 1919. For several years he spent each summer at the Mount Wilson Observatory, where he was appointed Research Assistant, and he generally visited Sir David Gill on the way out and back. Many a delightful evening was spent with a small circle of friends at Gill's house, when the problems of the universe were debated and plans for future work were laid. His good fellowship and enthusiasm and the simplicity of his character endeared him to his colleagues. His domestic life was singularly happy. Released from the duties of his official position early in 1921, he eagerly sought the renewed intercourse with his colleagues, which the lifting of the shadows of war made possible. He attended the meeting of the Astronomische Gesellschaft at Potsdam in August. At that meeting a little company of fellow-workers from Denmark, Sweden, Holland, England and Germany gathered together at Einstein's house to hear and debate his scheme of structure of the stellar universe; and we rejoiced to hear again the familiar guttural exclamations and quaint expressions, as with youthful spirit and enthusiasm he unfolded his latest ideas. A few days later he was at the British Association at Edinburgh, delighted to meet again the many friends of former years. Our last message from him was on the occasion of the Centenary of the Royal Astronomical Society. It was then known that a grave illness was beginning to develop, though he himself was full of hope of ultimate recovery until two days before the end. He died on June 18, 1922. By his death we lose one who has left a great mark upon the progress of astronomy, and has inspired many fellow-workers to follow up and extend his discoveries.

A. S. E.

INDEX TO VOL. CII. (A)

Absorption bands, wave-length measurement of (Hartridge), 575.
Address of the President, 1922, 373.
α- and β-rays, theory of scattering (Jeans), 437.
α-particle passing through matter, changes in charge of (Henderson), 496.
α-ray, loss of energy of, in passage through matter (Kapitza), 48.
α-ray photographs, analysis of (Blackett), 294.
Aluminium crystal, distortion during tensile test (Taylor and Elam), 643.
Argon, ionisation by electron collisions (Horton and Davies), 131.
Armstrong (E. F.) and Hilditch (T. P.) A Study of Catalytic Action at Solid Surfaces.— Part VIII, 21; Part IX, 27.
Astbury (W. T.) The Crystalline Structure and Properties of Tartaric Acid, 506.
Atmosphere, outer, and theory of meteors (Lindemann and Dobson), 411.
Atmospheric disturbances, directional observations of (Watt), 460.
Atomic hydrogen, incandescence in (Wood), 1.

Backhurst (I.) Variation of the Intensity of Reflected X radiation with the Temperature of the Crystal, 340.
Bakerian Lecture (Taylor and Elam), 643.
Berry (A.) and Swain (L. M.) On the Steady Motion of a Cylinder through Infinite Viscous Fluid, 766.
β-rays, scattering of (Wilson), 9.
Blackett (P. M. S.) On the Analysis of α-Ray Photographs, 294.
Brodetsky (S.) The Line of Action of the Resultant Pressure in Discontinuous Fluid Motion, 361; —— Discontinuous Fluid Motion past Circular and Elliptic Cylinders 542.

Cale (F. M.) See McLennan and Cale.
Campbell (J. H. P.) See Whytlaw-Gray.
Carbon arc ultra-violet spectrum (Simeon), 484.
Catalytic Action (Armstrong and Hilditch), 21, 27.
Christie (Sir W.) Obituary notice of, xi.
Clark (M. L.) See McLennan and Clark.
Colloidal behaviour, theory of (Hill), 705.
Colloids, periodic opacity of (Holker), 710.
Colour vision, investigation by Rayleigh test (Houstoun), 353.
Copper films, structure and chemical activity of (Hinshelwood). 318.
Cyanogen spectrum, note on origin of (Rayleigh), 453.

Davies (A. C.) See Horton and Davies.
Davy (N.) See Shaw and Davy.
Dobson (G. M. B.) See Lindemann and Dobson.
Donnan equilibrium and colloidal theory (Hill), 705.

Earthquake wave propagation (Jeans), 554.
Eddington (A. S.) The Propagation of Gravitational Waves, 268.

Elam (C. F.) See Taylor (G. I.) and Elam (C. F.).
Electric discharge, evolution of heat and supply of energy during (Roberts), 72.
Electrode potential drop, measurements of, with electrolysis (Marsh and Evans), 328.
Electrode, potential, significance of (Heyrovský), 628.
Electrolytes, opacity of colloids in (Holker), 710.
Electrons from tungsten filament, energy of (Jones), 734.
Evans (A. E.) See Marsh and Evans.

Fluid motion, discontinuous (Brodetsky), 361, 542.
Fluid, motion of cylinder through (Berry and Swain), 766 ; of particles in (Jeffery), 161.
Fluorescence of æsculin (McLennan and Cale), 256.

Gowland (W.) Obituary notice of, xvi.
Grace (S. F.) Free Motion of a Sphere in a Rotating Liquid parallel to the Axis of Rotation, 89.
Gravitational waves, propagation of (Eddington), 268.
Gravitative attraction, effect of temperature on (Shaw), 46.
Grove-Hills (E. H.) Obituary notice of, xx.
Gyromagnetic ratio, magnitude of (Richardson), 538.

Haloes, pleochroic, of various geological ages (Joly), 682.
Hartridge (H.) The Coincidence Method for the Wave-length Measurement of Absorption Bands, 575.
Henderson (G. H.) Changes in the Charge of an α-Particle passing through Matter, 496.
Heyrovský (J.) The Significance of the Electrode Potential, 628.
Hicks (W. M.) On the Mutual Threading of Vortex Rings, 111.
Hilditch (T. P.) See Armstrong and Hilditch.
Hill (A. V.) The Potential Difference occurring in a Donnan Equilibrium and the Theory of Colloidal Behaviour, 705.
Hinshelwood (C. N.) On the Structure and Chemical Activity of Copper Films, 318.
Holker (J.) The Periodic Opacity of Certain Colloids in progressively increasing Concentrations of Electrolytes, 710.
Horton (F.) and Davies (A. C.) A Spectroscopic Investigation of the Ionisation of Argon by Electron Collisions, 131.
Houstoun (R. A.) An Investigation of the Colour Vision of 527 Students by the Rayleigh Test, 353.
Hydrogen, incandescence in atomic (Wood), 1.
Hydrogen positive rays and field of force (Thompson), 197.
Hydrogenation of phenol (Armstrong and Hilditch), 21.

Ionisation of argon by electron collisions (Horton and Davies), 131.
Ionising potentials, new method for studying (Smyth), 283.
Iridescent colour and structure (Rayleigh), 668, 674.
Iron and steel at temperatures below 280° C. (Thompson and Whitehead), 587.

Jackson (L. C.) and Onnes (H. K.) Investigations on the Paramagnetic Sulphates at Low Temperatures, 678 ; —— —— Investigations on the Paramagnetism of Crystals at Low Temperatures, 680.
Jeans (J. H.) The Theory of the Scattering of α- and β-rays, 437 ; —— The Propagation of Earthquake Waves, 554.
Jeffery (G. B.) The Motion of Ellipsoidal Particles immersed in a Viscous Fluid, 161.
Joly (J.) Pleochroic Haloes of Various Geological Ages, 682.

Jones (J. H.) The Kinetic Energy of Electrons emitted from a Hot Tungsten Filament, 734.

Kapitza (P. L.) The Loss of Energy of an α-Ray Beam in its Passage through Matter. Part I.—Air and CO_2, 48.
Kapteyn (J. C.) Obituary notice of, xxix.
Kempe (Sir A. B.) Obituary notice of, i.
Knott (C. G.) Obituary notice of, xxvii.

Light, molecular scattering in vapours and liquids (Ramanathan), 151.
Light, scattered by mercury vapour, polarisation of (Rayleigh), 190.
Lindemann (F. A.) and Dobson (G. M. B.) A Theory of Meteors, and the Density and Temperature of the Outer Atmosphere to which it leads, 411.
Liquid, rotating (Grace), 89 ; (Taylor), 180.
Lubrication, cylindrical journal, at high eccentricity (Stanton), 241.

McAulay (A.) Multenions and Differential Invariants.—II, 210.
McLennan (J. C.) and others. On the Absorption of λ 5460·97 A. by Luminous Mercury Vapour, 33.
McLennan (J. C.) and Cale (F. M.) On the Fluorescence of Æsculin, 256.
McLennan (J. C.) and Clark (M. L.) On the Excitation of Characteristic X-Rays from Light Elements, 389.
Marsh (S.) and Evans (A. E.) On Measurements of Electrode Potential Drop with Direct and Alternating Current Electrolysis, 328.
Mercury vapour, absorption of λ 5460·97 A. by (McLennan and others), 33.
Meteors, theory of (Lindemann and Dobson), 411.
Mosharrafa (A. M.) On the Quantum Theory of the Simple Zeeman Effect, 529.
Multenions and invariants (McAulay), 210.

Nitrogen, spectrum of active (Rayleigh), 453.

Obituary Notices of Fellows deceased :—
Kempe (A. B.), i.
Christie (W. H. M.), xi.
Gowland (W.), xvi.
Grove-Hills (E. H.), xx.
Walker (G. W.), xxii.
Knott (C. G), xxvii.
Kapteyn (J. C.), xxix.

Onnes (H. K.) See Jackson and Onnes.

Paget (Sir R. A. S.) The Production of Artificial Vowel Sounds, 752.
Paramagnetic sulphates and crystals at low temperatures (Jackson and Onnes), 678, 680.

Quantum theory and electro-magnetic phenomena (Wilson), 478.
Quartz, effect of long grinding on (Ray), 640.

Ramanathan (K. R.) The Molecular Scattering of Light in Vapours and in Liquids and its Relation to the Opalescence observed in the Critical State, 151.
Ray (R. C.) The Effect of Long Grinding on Quartz (Silver Sand), 640.
Rayleigh (Lord) Polarisation of the Light Scattered by Mercury Vapour near the Resonance Periodicity, 190; —— Spectrum of Active Nitrogen as Affected by Admixture of the Inert Gases, with Note on Origin of Cyanogen Spectrum, 453; —— Studies of Iridescent Colour, and the Structure producing it.—I. Colours of Potassium Chlorate Crystals, 668 ; II. Mother-of-Pearl, 674.

Richardson (O. W.) The Magnitude of the Gyromagnetic Ratio, 538.
Roberts (J. K.) The Relation between the Evolution of Heat and the Supply of Energy during the Passage of an Electric Discharge through Hydrogen, 72.
Rotating liquid, motion of sphere in (Grace), 89 ; (Taylor), 180.

Shaw (P. E.) and Davy (N.) The Effect of Temperature on Gravitative Attraction, 46
Sherrington (Sir C. S.) Presidential Address, 1922, 373.
Simeon (F.) The Carbon Arc Spectrum in the Extreme Ultra-violet, 484.
Smokes, behaviour, number and size of particles (Whytlaw-Gray and others), 600.
Smyth (H. D.) A New Method for Studying Ionising Potentials, 283.
Speakman (J. B.) See Whytlaw-Gray.
Sphere in rotating liquid, motion of (Grace), 89 ; (Taylor), 180.
Stanton (T. E.) On the Characteristics of Cylindrical Journal Lubrication at High Values of the Eccentricity, 241.
Swain (L. M.) See Berry and Swain.

Tartaric acid, crystalline structure and properties (Astbury), 506.
Taylor (G. I.) The Motion of a Sphere in a Rotating Liquid, 180 ; —— Stability of a Viscous Liquid contained between Two Rotating Cylinders, 541.
Taylor (G. I.) and Elam (C. F.) The Distortion of an Aluminium Crystal during a Tensile Test, 643.
Tensile distortion of aluminium (Taylor and Elam), 643.
Thompson (F. C.) and Whitehead (E.) On the Changes in Iron and Steel at Temperatures below 280° C., 587.
Thomson (G. P.) The Scattering of Hydrogen Positive Rays, and the Existence of a Powerful Field of Force in the H Molecule, 197.

Ultra-violet carbon arc spectrum (Simeon), 484.

Viscous fluid, motion of particles in (Jeffery), 161.
Vortex rays, mutual threading of (Hicks), 111.
Vowel sounds, production of artificial (Paget), 752.

Walker (G. W.) Obituary notice of, xxii.
Watt (R. A. W.) Directional Observations of Atmospheric Disturbances, 1920-21, 460.
Whitehead (E.) See Thompson and Whitehead.
Whytlaw-Gray (R. and others) Smokes. Part I.—A Study of Behaviour and Number of Particles, 600. Part II.—Size of the Particles, 615.
Wilson (H. A.) On the Scattering of β-Rays, 9.
Wilson (W.) The Quantum Theory and Electromagnetic Phenomena, 478.
Wood (R. W.) Spontaneous Incandescence of Substances in Atomic Hydrogen Gas, 1.

X-radiation, variation of the intensity of reflected (Backhurst), 340.
X-rays, excitation of, from light elements (McLennan and Clark), 389.

Zeeman effect, quantum theory of (Mosharrafa), 529.

END OF THE HUNDRED AND SECOND VOLUME (SERIES A).

I. A. R. I. 75

IMPERIAL AGRICULTURAL RESEARCH
INSTITUTE LIBRARY
NEW DELHI.

Date of issue.	Date of issue.	Date of issue.